

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 10:59:54 ; Search time 1711.61 Seconds
(without alignments)
192.768 Million cell updates/sec

Title: US-09-885-799-318

Perfect score: 1 cactgaagtaactaagaag 20

Sequence:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.om:*
20: em.ov:*
21: em.pat:*
22: em.ph:*
23: em.pl:*
24: em.ro:*
25: em.sts:*
26: em.sy:*
27: em.un:*
28: em.vl:*
29: em.fun:*
30: em.hum:*
31: em.in:*
32: em.om:*
33: em.ov:*
34: em.pat:*
35: em.ph:*
36: em.pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	22	6	114275	114275 Sequence 72
2	20	100.0	22	6	122732	122732 Sequence 22
3	20	100.0	22	6	147557	147557 Sequence 22
4	20	100.0	30	6	A46153	A46153 Sequence 48
5	20	100.0	449	14	HP045924	HP045924 Human papill
6	20	100.0	449	14	HP045925	HP045925 Human papill
7	20	100.0	449	14	HP045926	HP045926 Human papill
8	20	100.0	449	14	HP045927	HP045927 Human papill
9	20	100.0	449	14	HP045928	HP045928 Human papill
10	20	100.0	7824	14	PPH58	PPH58 Human papill
11	18.4	92.0	449	14	HP045929	HP045929 Human papill
12	18	90.0	176549	2	AC034139	AC034139 Homo sapi
13	18	90.0	185440	2	AC015499	AC015499 Homo sapi
14	17.4	87.0	580	11	G58130	G58130 SHGC-103964
15	17.4	87.0	114956	9	AC002072	AC002072 Human PAC
16	17.4	87.0	169118	2	AC073313	AC073313 Homo sapi
17	17.4	87.0	175645	9	AC073114	AC073114 Homo sapi
18	17.4	87.0	190912	9	AL357559	AL357559 Human DNA
19	17	85.0	120039	2	AL357124	AL357124 Homo sapi
20	17	85.0	155889	2	AC025642	AC025642 Homo sapi
21	17	85.0	157023	2	AL158171	AL158171 Homo sapi
22	17	85.0	162347	9	AL356108	AL356108 Human DNA
23	16.8	84.0	566	3	BM068169	BM068169 Brugia mala
24	16.8	84.0	1203	3	EPHER1AD	EPHER1AD Homo sapi
25	16.8	84.0	2483	8	AF152096	AF152096 Arabidops
26	16.8	84.0	6640	9	D63997	D63997 Homo sapien
27	16.8	84.0	32998	9	AP000322	AP000322 Homo sapi
28	16.8	84.0	63807	2	AC087299	AC087299 Homo sapi
29	16.8	84.0	96180	9	AC005868	AC005868 Homo sapi
30	16.8	84.0	100000	9	AP000053	AP000053 Homo sapi
31	16.8	84.0	100000	9	AP000121	AP000121 Homo sapi
32	16.8	84.0	100000	9	AP000167	AP000167 Homo sapi
33	16.8	84.0	100828	9	AL590558	AL590558 Human DNA
34	16.8	44.0	103560	8	F5114	AC001229 Sequence
35	16.8	84.0	112328	9	AL390997	AL390997 Human DNA
36	16.8	84.0	130114	2	AF127414	AF127414 Homo sapi
37	16.8	84.0	134787	2	AL356692	AL356692 Homo sapi
38	16.8	84.0	135648	9	AC004069	AC004069 Homo sapi
39	16.8	84.0	141025	2	AC091795	AC091795 Felis cat
40	16.8	84.0	147891	2	AL365191	AL365191 Homo sapi
41	16.8	84.0	153929	2	AC026023	AC026023 Homo sapi
42	16.8	84.0	162680	9	AL158203	AL158203 Human DNA
43	16.8	84.0	165742	2	AC020697	AC020697 Homo sapi
44	16.8	84.0	168925	2	AL591179	AL591179 Homo sapi
45	16.8	84.0	169364	3	CEY56A3A	AL132860 Caenorhab

ALIGNMENTS

RESULT 1
114275
LOCUS 114275 22 bp DNA
DEFINITION Sequence 72 from patent US 5447839.
ACCESSION 114275
VERSION 114275.1 GI:997290
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Manos/Michele, Bauer/H.M., Greer/C.E., Resnick/R.M. and Ting/Y.
TITLE Detection of human papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5447839-A 72 05-SPP-1995;
FEATURES
source location/Qualifiers
1..22
BASE COUNT 9 a 3 c 7 g 3 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 cactgaagtaactaagaag 20
 |||||||
 Db 2 CACTGAGTACTAAGAAG 21

RESULT 2
 LOCUS 122732
 DEFINITION Sequence 220 from patent US 5527898.
 ACCESSION 122732
 VERSION 122732.1 GI:1603086
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 22)
 AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Manos,M.Michele, Resnick,R.M. and Zhang,T.Y.
 TITLE Detection of human papillomavirus by the polymerase chain reaction
 JOURNAL Patent: US 5527898-A 220 18-JUN-1996;
 FEATURES Location/Qualifiers
 source 1..22
 /organism="unknown"

BASE COUNT 9 a 3 c 7 g 3 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 cactgaagtaactaagaag 20
 |||||||
 Db 2 CACTGAGTACTAAGAAG 21

RESULT 3
 LOCUS 147557
 DEFINITION Sequence 220 from patent US 5639871.
 ACCESSION 147557
 VERSION 147557.1 GI:2471522
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 22)
 AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Impraim,C.C., Manos,M.Michele, Resnick,R.M. and Zhang,T.Y.
 TITLE Detection of human papillomavirus by the polymerase chain reaction
 JOURNAL Patent: US 5639871-A 220 17-JUN-1997;
 FEATURES Location/Qualifiers
 source 1..22
 /organism="unknown"

BASE COUNT 9 a 3 c 7 g 3 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 cactgaagtaactaagaag 20
 |||||||
 Db 2 CACTGAGTACTAAGAAG 21

RESULT 4

A46153
 LOCUS A46153 30 bp DNA
 DEFINITION Sequence 48 from Patent WO9522626.
 ACCESSION A46153
 VERSION A46153.1 GI:2300401
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.

REFERENCE 1 (bases 1 to 30)
 AUTHORS Meijer,C.J., Van,D.B., Walboomers,J.M. and Snijders,P.J.
 TITLE HUMAN PAPILLOMA VIRUS DETECTION IN A NOCLEIC ACID AMPLIFICATION
 JOURNAL PROCESS USING GENERAL PRIMERS
 COMMENT Patent: WO 9522626-A 48 24-AUG-1995;
 FEATURES STICHTING RESEARCHFONDS PATROL (NI);
 OTHER publication AU 1672295 950904.
 LOCATION/Qualifiers
 source 1..30
 /organism="unidentified"
 /db_xref="taxon:3264"

BASE COUNT 12 a 4 c 7 g 7 t
 ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 30;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 cactgaagtaactaagaag 20
 |||||||
 Db 7 CACTGAGTACTAAGAAG 26

RESULT 5
 LOCUS HP045924 449 bp DNA
 DEFINITION Human papillomavirus type 58, isolate IS068, major capsid protein
 ACCESSION LI (LI) gene, partial cds.
 VERSION 045924
 KEYWORDS 045924.1 GI:1209174
 SOURCE Human papillomavirus type 58.
 ORGANISM Human papillomavirus type 58.

REFERENCE 1 (bases 1 to 449)
 AUTHORS Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X., Petro,J. and Wheeler,C.M.
 TITLE Initiatype variation in 12 human papillomavirus types: a worldwide
 JOURNAL perspective
 MEDLINE J. Virol. 70 (5), 3127-3136 (1996)
 REFERENCE 96186743
 AUTHORS 2 (bases 1 to 449)
 JOURNAL Farmer,A.D.

COMMENT Direct Submission
 Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
 Laboratory (HPV Sequence Database, T-10, Mail Stop K710, Los
 Alamos, NM 87545, USA
 IS068 was originally isolated from a cervical cancer specimen from
 Brazil, in connection with the study published in Bosch, et al., J.
 Natl. Cancer Inst. 87:796-802. A variant of HPV58, this isolate has
 been sequenced over the My09/11 consensus primer region, beginning
 658 bp downstream of the origin for HPV58. Due to primer
 degeneracy, sequence information over the primer regions is
 uncertain.

FEATURES
 source 1..449
 /organism="Human papillomavirus type 58"

/isolate="IS068"
 /db_xref="taxon:10598"
 1..20
 /gene="LI"
 primer_bind
 /note="My11 primer"

CDS

<1..>449
/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AA19159.1"
/db_xref="GI:1209175"
/translation="AAGHNGICWGNQLFVTVDTRSTMTLCTEVRKEGYKDNF
KEYVHVEEDLOFVQLCKITLTAEMTYIHTMDSNILEDQFGLTPPSASLQDTY
RFVTSOAITCOKTAPPEKEDPLNKYTFMEVNLKEKFSADLDQFPLGR"
1..449
/gene="L1"
/note="My09 primer"
/db_xref="taxon:10598"
1..20
/gene="L1"
/note="My11 primer"
/db_xref="taxon:10598"
1..20
/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AA19160.1"

FEATURES

source
1..449
/organism="Human papillomavirus type 58"
/isolate="IS573"
/db_xref="taxon:10598"
1..20
/gene="L1"
/note="My11 primer"
/db_xref="taxon:10598"
1..20
/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AA19160.1"

Query Match 100.0%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
|||||
Db 90 CACTGAAGTAACTAAGAG 109

RESULT 6
HPV45925 449 bp DNA VRL 20-NOV-1996
LOCUS Human papillomavirus type 58, isolate IS573, major capsid protein
DEFINITION L1 (L1) gene, partial cds.
ACCESSION U45925
VERSION U45925.1 GI:1209176
KEYWORDS
SOURCE Human papillomavirus type 58.
ORGANISM Human papillomavirus type 58.
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE 1 (bases 1 to 449)
AUTHORS Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.
TITLE Intratype variation in 12 human papillomavirus types: a worldwide
perspective
JOURNAL J. Virol. 70 (5), 3127-3136 (1996)
MEDLINE 96186743
REFERENCE 2 (bases 1 to 449)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, HPV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA
COMMENT IS573 was originally isolated from a cervical cancer specimen from
Paraguay, in connection with the study published in Bosch, et al.,
J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this isolate
has been sequenced over the My09/11 consensus primer region,
beginning 6588 bp downstream of the origin for HPV58. Due to primer
degeneracy, sequence information over the primer regions is
uncertain.

gene
/db_xref="GI:1209177"
/translation="AAGHNGICWGNQLFVTVDTRSTMTLCTEVRKEGYKDNF
KEYVHVEEDLOFVQLCKITLTAEMTYIHTMDSNILEDQFGLTPPSASLQDTY
RFVTSOAITCOKTAPPEKEDPLNKYTFMEVNLKEKFSADLDQFPLGR"
1..449
/gene="L1"
/note="My09 primer"
/db_xref="taxon:10598"
1..20
/gene="L1"
/note="My11 primer"
/db_xref="taxon:10598"
1..20
/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AA19161.1"
/db_xref="GI:1209179"
/translation="AAGHNGICWGNQLFVTVDTRSTMTLCTEVRKEGYKDNF
KEYVHVEEDLOFVQLCKITLTAEMTYIHTMDSNILEDQFGLTPPSASLQDTY
RFVTSOAITCOKTAPPEKEDPLNKYTFMEVNLKEKFSADLDQFPLGR"
1..449
/gene="L1"

FEATURES

source
1..449
/organism="Human papillomavirus type 58"
/isolate="IS1021"
/db_xref="taxon:10598"
1..20
/gene="L1"
/note="My11 primer"
/db_xref="taxon:10598"
1..20
/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AA19161.1"
/db_xref="GI:1209179"
/translation="AAGHNGICWGNQLFVTVDTRSTMTLCTEVRKEGYKDNF
KEYVHVEEDLOFVQLCKITLTAEMTYIHTMDSNILEDQFGLTPPSASLQDTY
RFVTSOAITCOKTAPPEKEDPLNKYTFMEVNLKEKFSADLDQFPLGR"
1..449
/gene="L1"

Query Match 100.0%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
|||||
Db 90 CACTGAAGTAACTAAGAG 109

RESULT 7
HPV45926 449 bp DNA VRL 20-NOV-1996
LOCUS Human papillomavirus type 58, isolate IS1021, major capsid protein
DEFINITION L1 (L1) gene, partial cds.
ACCESSION U45926
VERSION U45926.1 GI:1209178
KEYWORDS
SOURCE Human papillomavirus type 58.
ORGANISM Human papillomavirus type 58.
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE 1 (bases 1 to 449)
AUTHORS Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.
TITLE Intratype variation in 12 human papillomavirus types: a worldwide
perspective
JOURNAL J. Virol. 70 (5), 3127-3136 (1996)
MEDLINE 96186743
REFERENCE 2 (bases 1 to 449)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, HPV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA
COMMENT IS1021 was originally isolated from a cervical cancer specimen from
the Philippines, in connection with the study published in Bosch, et
al., J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this
isolate has been sequenced over the My09/11 consensus primer
region, beginning 6588 bp downstream of the origin for HPV58. Due
to primer degeneracy, sequence information over the primer regions
is uncertain.

primer_bind complement(429..449)

BASE COUNT 148 a 78 c 86 g 137 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
|||||
Db 90 CACTGAAGTACTAAGCAG 109

HPV45927 449 bp DNA VRL 20-NOV-1996
LOCUS Human papillomavirus type 58, isolate IS131, major capsid protein

DEFINITION L1 (L1) gene, partial cds.

ACCESSION U45927.1 GI:1209180

VERSION 1
KEYWORDS Human papillomavirus type 58.
SOURCE Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
ORGANISM

REFERENCE 1 (bases 1 to 449)
Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.
Intratype variation in 12 human papillomavirus types: a worldwide
perspective
J. Virol. 70 (5), 3127-3136 (1996)

JOURNAL MEDLINE
REFERENCE 2 (bases 1 to 449)
Farmer,A.D.
Direct Submission
Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, HPV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA

COMMENT IS131 was originally isolated from a cervical cancer specimen from
Bolivia, in connection with the study published in Bosch, et al.,
J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this isolate
has been sequenced over the My09/11 consensus primer region,
beginning 6588 bp downstream of the origin for HPV58. Due to primer
degeneracy, sequence information over the primer regions is
uncertain.

FEATURES

source

1..449
/organism="Human papillomavirus type 58"
/isolate="IS131"
/db_xref="taxon:10598"

primer_bind

1..20
/gene="L1"
/note="My11 primer"
<1..>449
/gene="L1"

CDS

/codon_start=1
/product="major capsid protein L1"
/protein_id="AAB19162.1"
/db_xref="GI:1209181"
/translation="AOGHNGICMGNOLEFVYVDTRSTJNMLCTEYTKETKYNDF
KEYVHVEYDLOFOLCKITLTAELMTYIHTMNSDILDMOGLTPPSASLDITY
RFVTSQAITCOKTAPPKEDPLNKITFEVNLKEKFSADLDQPLGR"
1..449
/gene="L1"
complement(429..449)
/note="My09 primer"

gene

primer_bind
BASE COUNT 150 a 78 c 84 g 137 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
|||||
Db 90 CACTGAAGTACTAAGCAG 109

HPV45928 449 bp DNA VRL 20-NOV-1996
LOCUS Human papillomavirus type 58, isolate IS404, major capsid protein

DEFINITION L1 (L1) gene, partial cds.

ACCESSION U45928.1 GI:1209182

VERSION 1
KEYWORDS Human papillomavirus type 58.
SOURCE Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
ORGANISM

REFERENCE 1 (bases 1 to 449)
Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.
Intratype variation in 12 human papillomavirus types: a worldwide
perspective
J. Virol. 70 (5), 3127-3136 (1996)

JOURNAL MEDLINE
REFERENCE 2 (bases 1 to 449)
Farmer,A.D.
Direct Submission
Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, HPV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA

COMMENT IS404 was originally isolated from a cervical cancer specimen from
Mali, in connection with the study published in Bosch, et al.,
J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this isolate has
been sequenced over the My09/11 consensus primer region, beginning
6588 bp downstream of the origin for HPV58. Due to primer
degeneracy, sequence information over the primer regions is
uncertain.

FEATURES

source

1..449
/organism="Human papillomavirus type 58"
/isolate="IS404"
/db_xref="taxon:10598"

primer_bind

1..20
/gene="L1"
/note="My11 primer"
<1..>449
/gene="L1"

CDS

/codon_start=1
/product="major capsid protein L1"
/protein_id="AAB19163.1"
/db_xref="GI:1209183"
/translation="AOGHNGICMGNOLEFVYVDTRSTJNMLCTEYTKETKYNDF
KEYVHVEYDLOFOLCKITLTAELMTYIHTMNSDILDMOGLTPPSASLDITY
RFVTSQAITCOKTAPPKEDPLNKITFEVNLKEKFSADLDQPLGR"
1..449
/gene="L1"
complement(429..449)
/note="My09 primer"

gene

primer_bind
BASE COUNT 149 a 78 c 85 g 137 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
|||||

Db 90 CACTGAGTAACTAGAGAG 109

RESULT 10

LOCUS PP58 7824 bp DNA circular VRL 25-JUL-1998

DEFINITION Human papillomavirus type 58 complete genome.

ACCESSION D90400

VERSION D90400.1 GI:222386

KEYWORDS

SOURCE Human papillomavirus type 58 (isolate:patient specimen GN479) DNA. Human papillomavirus type 58 viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.

REFERENCE 1 (bases 1 to 7824)
Kiril, Y., Iwamoto, S. and Matsukura, T.
Human papillomavirus type 58 DNA sequence
Virology 185 (1), 424-427 (1991)

JOURNAL 92024102

MEDLINE

COMMENT Data kindly submitted in computer readable form by: Yasuyuki Kiril
Kanebo Institute for Cancer Research
1-5-90 Tomodachi-cho
Miyakojima-Ku, Osaka 534
Japan
Phone: 06-921-1281
Fax: 06-921-3883

FEATURES

source location/Qualifiers

1..7824

/organism="Human papillomavirus type 58"

/isolate="patient specimen GN479"

/db_xref="taxon:10598"

/note="HPV58"

72..77

110..559

/note="open reading frame E6"

/codon_start=1

/protein_id="BAA31845.1"

/db_xref="GI:3337098"

/translation="MFDAEKKPRLHDLQALETSVAHLEIKVCEKKTKLORSEYD
FVADLRIVRDGNFVACVGLRLSKISEYRHNYSLVGLTEQLKGLNEILIR
CLICORLCPCKRHVDLNRPHNISGNWTCGACVCMRRKQY"

574..870

/note="open reading frame E7"

/codon_start=1

/protein_id="BAA31846.1"

/db_xref="GI:3337099"

/translation="MRGNPRLREYILDLHPEPTDLFCYEQLCDSEDEIGLDPDG
QAQPATANYIVTCYTCCTVRLCINSTTDVRLQLGLMGTCIVCPSCAQ"

883..2817

/note="open reading frame E1"

/codon_start=1

/protein_id="BAA31847.1"

/db_xref="GI:3337100"

/translation="MDDEGTNGVAGCTGMEVEAVIERRTGDNISDEDEETADNG
TDLIEFIDSVOSTQAEEAARALFNVOGVDINAVCAIKRFPACSSAEDCDVD
RAANVCWKKYKNECTHRRKRIIEEDSGYNTVEYRQAHVYEQNDALDNSE
SSGVCASSVSETDVDSCTNVPQINISLNHSNKATLLYKTKKAYGVSEMLVP
EKSDKTSCTDWCITGYSISVAESLKVILKHSYTHLQCLTDGRIILLLIRKCK
SKNRLIVAKMSNLISIPETCMIIPEPKLRSAQALYFRPAMSNISDVGGTPEMD
RLVTLQSHENDIDLSMIQMAVNDIDDSIDIAKYAQLADVNSNAAFELRNAQA
KIYKCGVACRRYKAEKRGKTMGOMIORSCEKTDGNNRPVIOFLRQNIETFAFL
VAKFQLOGVPRKSCMLCGPANTGSGYSGMLIFELGCLISVYNSKSHFQPLSD
AKGMIDDTATSWTYIDDKYRNALDGNIDSTDKHRLVQLKCPULITISNAGKD
SRKPYLHSLTYFERNNPFPDANGNPVKINDENWKSFSRTWCKLGLIEEDKEND
GGNISTFKCSAQGNRIHS"

2753..3829

/note="open reading frame E2"

/codon_start=1

/protein_id="BAA31848.1"

/db_xref="GI:3337101"

/translation="MEETISRLSAVQOKILDIYEADKNDLTSQIEHMKLIRMECAIM
TARQMGISHCHQVPSIVASKRAFOVIELMALETENASPYKDEMTLQOQTSLEW
LSBPCKCFKKKGIYVQYNDKANIMDYTNSEIYIIEETCTLVAGEVDYGLYI

CDS

CGTGAAGTAACTAGAGAG 6696

Query Match 100.0%; Score 20; DB 14; Length 7824;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6677 CACTGAGTAACTAGAGAG 6696

RESULT 11

LOCUS HP045929 449 bp DNA VRL 20-NOV-1996

DEFINITION Human papillomavirus type 58, isolate IS417, major capsid protein L1 (L1) gene, partial cds.

ACCESSION U45929

VERSION U45929.1 GI:1209184

KEYWORDS

SOURCE Human papillomavirus type 58. Human papillomavirus type 58 viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.

REFERENCE 1 (bases 1 to 449)

CDS

2487 a 1388 c 1576 g 2373 t

BASE COUNT

ORIGIN

100.0%; Score 20; DB 14; Length 7824;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 20; DB 14; Length 7824;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6677 CACTGAGTAACTAGAGAG 6696

RESULT 11

LOCUS HP045929 449 bp DNA VRL 20-NOV-1996

DEFINITION Human papillomavirus type 58, isolate IS417, major capsid protein L1 (L1) gene, partial cds.

ACCESSION U45929

VERSION U45929.1 GI:1209184

KEYWORDS

SOURCE Human papillomavirus type 58. Human papillomavirus type 58 viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.

REFERENCE 1 (bases 1 to 449)

AUTHORS Stewart, A.C., Eriksson, A.M., Manos, M.M., Munoz, N., Bosch, F.X., Peto, J., and Wheeler, C.M.
 TITLE Initiate variation in 12 human papillomavirus types: a worldwide perspective
 JOURNAL U. Virol. 70 (5), 3127-3136 (1996)
 MEDLINE 96186743
 REFERENCE 2 (bases 1 to 449)
 AUTHORS Farmer, A.D.
 TITLE Direct submission
 JOURNAL Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National Laboratory, HPV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87545, USA
 COMMENT 15417 was originally isolated from a cervical cancer specimen from Mail, in connection with the study published in Bosch, et al., J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this isolate has been sequenced over the My09/11 consensus primer region, beginning 6588 bp downstream of the origin for HPV58. Due to primer degeneracy, sequence information over the primer regions is uncertain.

FEATURES
 source location/Qualifiers
 1..449
 /organism="Human papillomavirus type 58"
 /isolate="15417"
 primer_bind /db_xref="taxon:10598"
 1..20
 /gene="L1"
 /note="My11 primer"
 <1..>449
 /gene="L1"
 /note="My09/11"
 /codon_start=1
 /product="major capsid protein L1"
 /protein_id="AA819164.1"
 /db_xref="GI:1209185"
 /translation="AAGHNNNGIOWGNOLEFVVDTRSTNNLTCTEYTKEDYKNNF
 KEYVHVSEYDQFQICLITLAEMVYIHTNSDIIEDMOGCLTPPSASIDDT
 RVTISQAITCCKTAPKEKEDPLMKTFEENLKKPSADIDPPLGR"
 1..449
 /gene="L1"
 complement(429..449)
 primer_bind /note="My09 primer"
 BASE COUNT 151 a 78 c 83 g 137 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 14; Length 449;
 Best Local Similarity 95.0%; Pred. NO. 81;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
 |||||
 Db 90 CACTGAGTACTAAGAAG 109

RESULT 12
 AC034139 176549 bp DNA HTG 19-AUG-2001
 LOCUS Homo sapiens chromosome 4 clone RP11-67M24, WORKING DRAFT SEQUENCE,
 ACCESSION AC034139
 VERSION 3 unrounded pieces.
 KEYWORDS AC034139.5 GI:15212504
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 176549)
 AUTHORS Waterston, R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 176549)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission

JOURNAL
 COMMENT Submitted (04-APR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 On Aug 19, 2001 this sequence version replaced g1:9739321.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Project Information
 Center project name: H_NH0067M24

----- Summary Statistics -----
 Sequencing vector: M13, 54%
 Sequencing vector: plasmid, 46%
 Chemistry: dye-terminator Big Dye, 46% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 175335 bases at least Q40
 Consensus quality: 175759 bases at least Q30
 Consensus quality: 175972 bases at least Q20
 Insert size: 194000; agarose-tp
 Insert size: 176349; sum-of-ctligs
 Quality coverage: 8.99 in Q20 bases; agarose-tp
 Quality coverage: 9.29 in Q20 bases; sum-of-ctligs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 16242: contig of 16242 bp in length
 * 16243 16342: gap of unknown length
 * 16343 69644: contig of 53302 bp in length
 * 69645 69744: gap of unknown length
 * 69745 176549: contig of 106805 bp in length.

FEATURES
 source location/Qualifiers
 1..176549
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RP11-67M24"
 1..16242
 /note="assembly_name:Contig24"
 16343..69644
 /note="assembly_name:Contig25
 clone_end:47
 vector_side:right"
 69745..176549
 /note="assembly_name:Contig26
 clone_end:SP6
 vector_side:right"
 ORIGIN

BASE COUNT 54747 a 36029 c 35469 g 50101 t 203 others

Query Match 90.0%; Score 18; DB 2; Length 176549;
 Best Local Similarity 100.0%; Pred. NO. 1;le+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 actgaagtaactaagaag 19
 |||||
 Db 139548 ACTGAGTACTAAGAA 139565

RESULT 13
 AC015499/c 185440 bp DNA HTG 08-NOV-2000
 LOCUS Homo sapiens clone RP11-21G18, WORKING DRAFT SEQUENCE, 36 unordered
 DEFINITION pieces.

ACCESSION AC015499
 VERSION AC015499.4 GI:10047766
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 185440)
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE Homo sapiens, clone RP11-21G18
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 185440)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
 Baldwin,J., Barina,N., Beckert,M., Boguslavsky,L., Boukigalter,B.,
 Brown,A., Castle,A., Colangelo,R., Collins,S., Collingmore,A.,
 Cooke,P., Deaellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
 Ferreira,P., Fitzhugh,M., Forrest,C., Funke,R., Gage,D.,
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karstad,A., Klein,J.,
 Lehoczy,J., Lieu,C., Locke,K., MacDonald,P., Marcus,N.,
 McEwan,P., McQuirk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 Peterson,K., Pollara,V., Riley,K., Roy,A., Santos,R., Severy,P.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 10, 2000 this sequence version replaced q1:7382153.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

Project Information
 Center project name: L4006

Center clone name: 21.C.18

Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 167969 bases at least Q40
 Consensus quality: 176547 bases at least Q30
 Consensus quality: 179506 bases at least Q20
 Insert size: 152000; agarose-tp
 Insert size: 181940; sum-of-ctnigs
 Quality coverage: 5.6 in Q20 bases; agarose-tp
 Quality coverage: 4.7 in Q20 bases; sum-of-ctnigs

NOTE: This is a 'working draft' sequence. It currently
 consists of 36 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 10462: contig of 10462 bp in length
 * 10463 10562: gap of 100 bp
 * 10563 11934: contig of 1372 bp in length
 * 11935 12034: gap of 100 bp
 * 12035 13303: contig of 1269 bp in length
 * 13304 13403: gap of 100 bp
 * 13404 14428: contig of 1025 bp in length
 * 14429 14528: gap of 100 bp
 * 14529 15938: contig of 1411 bp in length
 * 15940 16039: gap of 100 bp
 * 16040 17356: contig of 1317 bp in length
 * 17357 17456: gap of 100 bp

FEATURES

Source

17457 18699: contig of 1243 bp in length
 * 18700 18799: gap of 100 bp
 * 18800 20052: contig of 1253 bp in length
 * 20053 20152: gap of 100 bp
 * 20153 21271: contig of 1119 bp in length
 * 21272 21371: gap of 100 bp
 * 21372 22382: contig of 1011 bp in length
 * 22383 22482: gap of 100 bp
 * 22483 23602: contig of 1120 bp in length
 * 23603 23702: gap of 100 bp
 * 23703 24959: contig of 1257 bp in length
 * 24960 25059: gap of 100 bp
 * 25060 26122: contig of 1063 bp in length
 * 26123 26222: gap of 100 bp
 * 26223 27314: contig of 1092 bp in length
 * 27315 27414: gap of 100 bp
 * 27415 29057: contig of 1643 bp in length
 * 29058 29157: gap of 100 bp
 * 29158 30171: contig of 1014 bp in length
 * 30172 30271: gap of 100 bp
 * 30272 31284: contig of 1013 bp in length
 * 31285 31384: gap of 100 bp
 * 31385 32423: contig of 1039 bp in length
 * 32424 32523: gap of 100 bp
 * 32524 33951: contig of 1428 bp in length
 * 33952 34051: gap of 100 bp
 * 34052 35667: contig of 1616 bp in length
 * 35668 35767: gap of 100 bp
 * 35768 37483: contig of 1716 bp in length
 * 37484 37583: gap of 100 bp
 * 37584 38894: contig of 1311 bp in length
 * 38895 38994: gap of 100 bp
 * 38995 40762: contig of 1768 bp in length
 * 40763 40862: gap of 100 bp
 * 40863 41953: contig of 1091 bp in length
 * 41954 42053: gap of 100 bp
 * 42054 43584: contig of 1531 bp in length
 * 43585 43684: gap of 100 bp
 * 43685 47261: contig of 3577 bp in length
 * 47262 47361: gap of 100 bp
 * 47362 50113: contig of 2752 bp in length
 * 50114 50213: gap of 100 bp
 * 50214 52333: contig of 2120 bp in length
 * 52334 52433: gap of 100 bp
 * 52434 56904: contig of 4471 bp in length
 * 56905 57004: gap of 100 bp
 * 57005 62798: contig of 5794 bp in length
 * 62799 62898: gap of 100 bp
 * 62899 68982: contig of 6084 bp in length
 * 68983 69082: gap of 100 bp
 * 69083 77983: contig of 8901 bp in length
 * 77984 78083: gap of 100 bp
 * 78084 97647: contig of 19564 bp in length
 * 97648 97747: gap of 100 bp
 * 97748 141022: contig of 43275 bp in length
 * 141023 141122: gap of 100 bp
 * 141123 165032: contig of 23910 bp in length
 * 165033 165132: gap of 100 bp
 * 165133 185440: contig of 20308 bp in length.

Location/Qualifiers

1. 185440
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-21G18"
 /clone_id="RP11-21 Human Male BAC"
 1. 10462
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 10563. 11934
 /note="assembly_fragment"
 12035. 13303
 /note="assembly_fragment"

```

misc_feature      13404..14428
                    /note="assembly_fragment"
misc_feature      14529..15939
                    /note="assembly_fragment"
misc_feature      16040..17356
                    /note="assembly_fragment"
misc_feature      17457..18699
                    /note="assembly_fragment"
misc_feature      18800..20052
                    /note="assembly_fragment"
misc_feature      20153..21271
                    /note="assembly_fragment"
misc_feature      21372..22382
                    /note="assembly_fragment"
misc_feature      22483..23602
                    /note="assembly_fragment"
misc_feature      23703..24959
                    /note="assembly_fragment"
misc_feature      25060..26122
                    /note="assembly_fragment"
misc_feature      26223..27314
                    /note="assembly_fragment"
misc_feature      27415..29057
                    /note="assembly_fragment"
misc_feature      29158..30171
                    /note="assembly_fragment"
misc_feature      30272..31284
                    /note="assembly_fragment"
misc_feature      31385..32423
                    /note="assembly_fragment"
misc_feature      32524..33951
                    /note="assembly_fragment"
misc_feature      34052..35667
                    /note="assembly_fragment"
misc_feature      35768..37483
                    /note="assembly_fragment"
misc_feature      37584..38894
                    /note="assembly_fragment"
misc_feature      38995..40762
                    /note="assembly_fragment"
misc_feature      40863..41953
                    /note="assembly_fragment"
misc_feature      42054..43584
                    /note="assembly_fragment"
misc_feature      43685..47261
                    /note="assembly_fragment"
misc_feature      47362..50113
                    /note="assembly_fragment"
misc_feature      50214..52333
                    /note="assembly_fragment"
misc_feature      52434..55904
                    /note="assembly_fragment"

Query Match      90.0%; Score 18; DB 2; Length 185440;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 actgaagtaactaagaa 19
        |||||
Db 106150 ACTGAGTAAGTAAGCA 106133

RESULT 14
LOCUS      G58130      580 bp      DNA
DEFINITION SHGC-103964 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION  G58130
VERSION    G58130.1 GI:6123299
KEYWORDS   STS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 580)

```

```

AUTHORS      Olivier, M. and Cox, D.R.
TITLE        Unpublished, Olivier, M., COX, D.R. (2000)
JOURNAL      Unpublished (2000)
COMMENT

Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: Olivier@shgc.stanford.edu
Primer A: TTGTCAGTGAAGTTTCCCTTT
Primer B: CCAGCAACACATTGTTCTCA
STS size: 293
PCR Profile:
  Initial incubation: 95 degrees C for 10 minutes
  Denaturation:      94 degrees C for 30 seconds
  Annealing:         60 degrees C for 30 seconds
  Polymerization:    72 degrees C for 23 seconds
  PCR cycles:        30
  Thermal Cycler:    Perkin Elmer 9700

Protocol:
  Template:          25 ng
  Primer:             each 1 uM
  dNTPs:             each 200 uM
  AmpliTaq Gold Polymerase: 0.07 units/ul
  Total Vol:         5 ul

Buffer:
  MgCl2:             2.5 mM
  KCl:               50 mM
  Tris-HCl:          10 mM
  pH:               8.3

FEATURES
  source
    1..580
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /map="1"
    /clone_lib="Human"
    STS
    primer_bind
      110..402
      primer_bind
        complement(380..402)
    BASE COUNT
      175 a 112 c 103 g 187 t 3 others
    ORIGIN

Query Match      87.0%; Score 17.4; DB 11; Length 580;
Best Local Similarity 94.7%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 actgaagtaactaagaa 20
        |||||
Db 279 ACTGAGTAAGTAAGAG 297

RESULT 15
LOCUS      AC002072/c 114956 bp      DNA
DEFINITION Human PAC clone RPI-218B13 from Xq23, complete sequence.
ACCESSION  AC002072
VERSION    AC002072.1 GI:2078473
KEYWORDS   HMG.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 114956)
AUTHORS      Graves, T. and Fronick, B.
TITLE        The sequence of H. sapiens PAC clone RPI-218B13
JOURNAL      Unpublished

```

REFERENCE 2 (bases 1 to 114956)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (12-MAY-1997)
 REFERENCE 3 (bases 1 to 114956)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 4 (bases 1 to 114956)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (04-FEB-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics -----
 Center project name: HLDJ218B13

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by David Bentley's chromosome X mapping group at the Sanger Centre. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX/>

SOURCE INFORMATION:
 This clone was derived from human PAC library RPCI-1 prepared by Pieter de Jong and coworkers at Roswell Park Cancer Institute, using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. For further details, see <http://bacpac.med.buffalo.edu/>. The clone is available from Genome Systems, Inc. (<http://www.genomesystems.com>).
 VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
 The actual start of this clone is at base position 1 of RPI-218B13; actual end is at 114956 of RPI-218B13. The orientation of this clone is unknown.

FEATURES
 This clone contains STR AFM184VA3 (NID:91051493).

SOURCE
 1. 114956
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /clone="RPI-218B13"
 /clone_lib="RPCI-1"
 /map="Xq23"
 /complement(66..385)
 repeat_region
 /rpt_family="L1"
 1236..1344
 /rpt_family="ALU"
 repeat_region
 /complement(1453..1745)
 /rpt_family="ALU"
 1763..1936
 /rpt_family="ALU"
 repeat_region
 5090..5379

repeat_region
 /rpt_family="ALU"
 5380..5424
 /rpt_family="L1"
 repeat_region
 /complement(5545..5843)
 /rpt_family="ALU"
 6476..6539
 /rpt_family="L1"
 repeat_region
 /complement(6613..6749)
 /rpt_family="ALU"
 7602..8081
 /rpt_family="L1"
 repeat_region
 8082..8375
 /rpt_family="ALU"
 8384..8533
 /rpt_family="L1"
 repeat_region
 10945..11332
 /rpt_family="L1"
 repeat_region
 11709..11733
 /rpt_family="L1"
 repeat_region
 /complement(14103..14393)
 /rpt_family="ALU"
 14961..15254
 /rpt_family="ALU"
 repeat_region
 /complement(15477..15765)
 /rpt_family="ALU"
 repeat_region
 /complement(16349..16642)
 /rpt_family="ALU"
 repeat_region
 /complement(16821..16840)
 /rpt_family="L1"
 repeat_region
 /complement(18232..18520)
 /rpt_family="ALU"
 repeat_region
 /complement(20661..21030)
 /rpt_family="MER"
 23813..24103
 /rpt_family="ALU"
 repeat_region
 /complement(27296..27590)
 /rpt_family="ALU"
 repeat_region
 /complement(28473..28512)
 /rpt_family="L1"
 28807..29089
 /rpt_family="ALU"
 repeat_region
 30181..30763
 /rpt_family="ALU"
 repeat_region
 30863..30884
 /rpt_family="ALU"
 repeat_region
 /complement(31105..31147)
 /rpt_family="L1"
 repeat_region
 /complement(31236..31527)
 /rpt_family="ALU"
 32649..32940
 /rpt_family="ALU"
 repeat_region
 33729..33764
 /rpt_family="ALU"
 repeat_region
 35472..35560
 /rpt_family="L1"
 repeat_region
 40761..41035
 /rpt_family="ALU"
 repeat_region
 /complement(42815..43105)
 /rpt_family="ALU"
 43245..43290
 /rpt_family="L1"
 repeat_region
 43434..43724
 /rpt_family="ALU"
 repeat_region
 /complement(45824..46132)
 /rpt_family="ALU"
 46450..46478
 /rpt_family="L1"
 repeat_region
 /complement(47883..49428)
 /rpt_family="L1"
 48971..49387
 /rpt_family="L1"
 repeat_region
 /complement(49511..50032)
 /rpt_family="L1"

```

repeat_region complement(50033 .50128)
/rpt_family="ALU"
repeat_region complement(50185 .50258)
/rpt_family="L1"
repeat_region complement(50259 .50375)
/rpt_family="ALU"
repeat_region complement(50867 .51148)
/rpt_family="ALU"
repeat_region complement(51149 .51192)
/rpt_family="L1"
repeat_region complement(51205 .51513)
/rpt_family="ALU"
repeat_region complement(52173 .52464)
/rpt_family="ALU"
repeat_region complement(52929 .52967)
/rpt_family="L1"
repeat_region complement(54069 .53779)
/rpt_family="ALU"
repeat_region complement(54296 .54165)
/rpt_family="L1"
repeat_region complement(58487 .58678)
/rpt_family="L1"
repeat_region complement(58683 .58971)
/rpt_family="ALU"
repeat_region complement(58972 .61254)
/rpt_family="L1"
repeat_region complement(60245 .59837)
/rpt_family="L1"
repeat_region complement(64958 .61621)
/rpt_family="L1"
repeat_region complement(63549 .63967)
/rpt_family="L1"
repeat_region complement(65023 .64992)
/rpt_family="MER"
repeat_region complement(65954 .66111)
/rpt_family="ALU"
repeat_region complement(67839 .67614)
/rpt_family="ALU"
repeat_region complement(67920 .67948)
/rpt_family="L1"
repeat_region complement(68209 .68250)

```

```

Query Match      87.0%; Score 17.4; DB 9; Length 114956;
Best Local Similarity 94.7%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 actgaactactagaag 20
Db 96472 ACTGAAGTAAATAGAG 96454

```

Search completed: April 3, 2002, 11:00:04
 Job time: 5195 sec

PT Detection of human papilloma

PT specific consensus primer pairs and pref. detection with generic or
 PT type specific probes for use in research and diagnosis
 XX
 PS Claim 3: Columns 61-62; 36pp; English.
 XX
 CC The human papilloma virus (HPV) specific probes AA098584-Q98650 are
 CC used to detect, or type HPV for research or diagnostic purposes,
 CC e.g. to identify HPV that are implicated in genital or oral
 CC carcinomas.
 XX
 SQ Sequence 22 BP; 9 A; 3 C; 7 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 16; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
 |||||
 Db 2 cactgaagtaactaagaag 21

RESULT 2
 AAT44681
 ID AAT44681 standard; DNA; 22 BP.

XX AC AAT44681;
 XX
 IT 28-JAN-1997 (first entry)
 XX

DE Human papillomavirus detection probe MY93 for HPV type 58.

KW Probe; primer; PCR; polymerase chain reaction; amplification;
 KM human papillomavirus; consensus; ss.

XX Synthetic.

XX US5527898-A.

XX 18-JUN-1996.

XX PF 09-SEP-1988; 88US-0243486.

XX PR 24-SEP-1993; 93US-0126452.

XX PR 09-SEP-1988; 88US-0243486.

XX PR 10-MAR-1989; 89US-0322550.

XX PR 09-SEP-1988; 89MO-US03747.

XX PR 14-NOV-1990; 90US-0613142.

XX PR 20-APR-1993; 93US-0050743.

XX PR 07-JUN-1995; 95US-0474542.

XX PA (HOFF) HOFFMANN I.A. ROCHE INC.

XX PI Bauer HM, Gravitt PE, Greer CE, Manos MM, Resnick RM;
 PI Zhang TY;

XX DR WPI; 1996-299903/30.

XX PT Nucleic acid hybridisation probes - specific for selected human
 PT papilloma virus types

XX PS Claim 1; Column 147; 96pp; English.

XX The invention relates to new oligonucleotide probes and primers used
 CC for the detection of human papillomaviruses which are not genital types
 CC 6, 11, 16, 18 or 33. The probes and primers AAT44608-T44693 are esp.
 CC used to detect HPV types 26, 31, 31B, 35, 39, 40, 43, 45, 51-59 and 68.

CC The primers can be used to detect these HPV types in conjunction with
 CC the consensus primers and typing probes AAT44733-T44906, which are based
 CC on and amplify fragments of the L1, E6, E7 and E1 regions of the HPV
 CC sequences. This primer is targeted to the new HPV type 58.

XX Sequence 22 BP; 9 A; 3 C; 7 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 17; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
 |||||
 Db 2 cactgaagtaactaagaag 21

RESULT 3
 AAT78084
 ID AAT78084 standard; DNA; 22 BP.

XX AC AAT78084;
 XX

XX 07-OCT-1997 (first entry)

DE Human papillomavirus 58 specific typing probe MY93.

KM Human; papillomavirus 58; HPV58; typing probe;
 KM detection; ss.

XX Synthetic.

XX US5639871-A.

XX 17-JUN-1997.

XX PF 09-SEP-1988; 88US-0243486.

XX PR 24-SEP-1993; 93US-0126452.

XX PR 09-SEP-1988; 88US-0243486.

XX PR 10-MAR-1989; 89US-0322550.

XX PR 09-SEP-1988; 89MO-US03747.

XX PR 14-NOV-1990; 90US-0613142.

XX PR 20-APR-1993; 93US-0050743.

XX PR 01-JUN-1995; 95US-0457648.

XX PA (HOFF) ROCHE MOLECULAR SYSTEMS INC.

XX PI Bauer HM, Gravitt PE, Greer CE, Imprim CC, Manos MM;
 XX Resnick RM, Zhang TY;

XX DR WPI; 1997-332084/30.

XX PT New oligonucleotide probes for human papilloma-virus - used for
 PT detecting and typing HPV and for detecting previously unknown HPV
 PT types and subtypes

XX PS Disclosure: Columns 145-146; 94pp; English.

XX CC The present sequence is a human papillomavirus 58 (HPV58)
 CC specific typing probe.

XX SQ Sequence 22 BP; 9 A; 3 C; 7 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 18; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaag 20
 |||||
 Db 2 cactgaagtaactaagaag 21

RESULT 4
 AAV17529
 ID AAV17529 standard; DNA; 22 BP.

XX AC AAV17529;


```

XX 04-JUN-1998 (first entry)
DT
XX
XX Probe MY93 for human papillomavirus typing.
DE
XX
XX Human papillomavirus: HPV; HPV detection; HPV typing;
KM L1 type-specific probe; ss.
KW
XX
XX Synthetic.
OS Human papillomavirus.
XX
XX US5705627-A.
PN
XX
XX 06-JAN-1998.
PD
XX
XX 26-MAY-1995; 95US-0452055.
PF
XX
XX 09-SEP-1989; 89US-0613142.
PR
XX 09-SEP-1988; 88US-0243486.
PR 10-MAR-1989; 89US-0322550.
PR 20-APR-1993; 93US-0050743.
PR 26-MAY-1995; 95US-0452055.
XX
XX (HOFF ) ROCHE MOLECULAR SYSTEMS INC.
XX
XX Bauer HM, Greer CE, Manos MM, Resnick RM, Ting Y;
PI WPI: 1998-192210/17.
DR
XX
XX Human papilloma probes and primers - useful for, e.g. detecting and
PT typing of human papilloma viruses
XX
XX Disclosure; Column 17-18: 37pp; English.
XX
XX This sequence represents a human papillomavirus (HPV) L1 type-specific
CC probe of the invention. This sequence may be used in conjunction with L1
CC specific primers for detecting and typing HPV. Identification and typing
CC of HPV is important as different types of HPV pose different risks for
CC infected individuals. HPV16 and HPV18 have been more consistently
CC identified in higher grades of cervical dysplasia and carcinoma than
CC other HPV types.
XX
XX Sequence 22 BP; 9 A; 3 C; 7 G; 3 T; 0 other;
SQ

```

Query Match 100.0%; Score 20; DB 19; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cactgaagtaactaagaag 20
    |||||||
Db 2 cactgaagtaactaagaag 21

```

RESULT 5
 AAT00997
 ID AAT00997 standard; DNA; 30 BP.
 XX
 XX AAT00997;
 AC
 XX
 XX 02-APR-1996 (first entry)
 DT
 XX
 XX Human papilloma virus HPV-58-specific probe.
 DE
 XX
 XX Human papilloma virus; HPV-58; cervical carcinoma; screening; PCR;
 KW elongated general primer; nucleic acid amplification; probe; ss.
 XX
 XX Synthetic.
 OS
 XX W09522626-A1.
 PN
 XX 24-AUG-1995.
 PD
 XX

```

PF 20-FEB-1995; 95WO-NL00066.
XX
XX 23-SEP-1994; 94EP-0202739.
PR 21-FEB-1994; 94EP-0200432.
XX
XX (RESE-) STICHTING RESEARCHFONDS PATHOLOGIE.
PA
XX
XX Meijer CJLM, Snijders PJF, Van Den Brule AJC, Walboomers JMM;
PI WPI: 1995-302728/39.
DR
XX
XX Elongated general primer(s) and Human Papilloma Virus-specific
PT probe(s) - for use in amplification and detection method(s)
PT providing improved HPV detection in cervical smear(s).
XX
XX Claim 29; Page 57; 61pp; English.
XX
XX Elongated general primers GP5+ and GP6+ and their substn. derivs.
CC are used for amplifying Human Papilloma Virus sequences. Unlike the
CC known general primers GP5 and GP6, originally selected from the HPV
CC L1 region on the basis of sequence information of HPV-6, -16, -18,
CC -31 and -33, the 3'-elongated primers overcome reduced PCR efficiency
CC and can distinguish HPV status in cytomorphologically normal cervical
CC smears which previously gave ambiguous or negative results.
CC The present sequence is that of a probe used for detecting HPV-58
CC sequences among PCR products amplified by the elongated primers.
XX
XX Sequence 30 BP; 12 A; 4 C; 7 G; 7 T; 0 other;
SQ

```

Query Match 100.0%; Score 20; DB 16; Length 30;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cactgaagtaactaagaag 20
    |||||||
Db 7 cactgaagtaactaagaag 26

```

RESULT 6
 AAV37221
 ID AAV37221 standard; DNA; 139 BP.
 XX
 XX AAV37221;
 AC
 XX
 XX 14-SEP-1998 (first entry)
 DT
 XX
 XX Target sequence in the L1 gene of HPV type 58.
 DE
 XX
 XX L1 gene; HPV type 58; hybridisation assay; detection; HPV type;
 KW cervical cancer; ss.
 XX
 XX Human papillomavirus.
 OS
 XX W09817829-A2.
 PN
 XX 30-APR-1998.
 PD
 XX
 XX 17-OCT-1997; 97WO-US19467.
 PF
 XX
 XX 25-OCT-1996; 96US-0739103.
 PR
 XX
 XX (ABBO) ABBOTT LAB.
 PA
 XX
 XX Abravaya K, Gorzowski JJ, Hoenle RJ, Kroeger PE;
 PI Moore JJ;
 PI WPI: 1998-261520/23.
 DR
 XX
 XX Detection of oncogenic human papilloma virus types - using
 PT oligonucleotide probes which hybridise with a 140 bp region of the
 PT L1 gene found in the human papillomavirus genome
 XX

PS Disclosure; Page 27; 33pp; English.

CC The present sequence represents a portion of the L1 gene of oncogenic
CC human papillomavirus (HPV) type 58. Oligonucleotide AAV38493 is used as
CC a primer and probe to detect this organism. The probes of the invention
CC can be used in a hybridisation assay to detect HPV types. They can also
CC be used in a similar way with the addition that HPV target sequences
CC can be amplified first. The products can be used for the detection of
CC HPV types, particularly those linked to the development of cervical
CC cancer.

SQ Sequence 139 BP; 47 A; 18 C; 26 G; 48 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cactgaagtaactaagaag 20
Db 48 cactgaagtaactaagaag 67

RESULT 7

AAT91255
ID AAT91255 standard; DNA; 449 BP.

AC AAT91255;

DT 24-APR-1998 (first entry)

DE Human papilloma virus HPV-58 DNA sequence for generating probes.

KM Human papilloma virus; HPV; probe; isolation; detection; hybridise; ds.

OS human papillomavirus.

PH Key Location/Qualifiers

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

XX Nucleic acid probes for human papilloma virus - for isolation and
PT detection of homologous or variant regions of DNA from different
PT viruses

PS Claim 8; Page 26; 48pp; English.

CC The present sequence represents a DNA sequence specific to a human
CC papilloma virus (HPV) which is used to generate specifically claimed
CC probes (see features table). The probes can be used to hybridise to
CC the homologous regions of the specific HPV. The probes can also be
CC used to generate probes with nucleotide substitutions, which will
CC hybridise to newly discovered variant regions of HPV. The probes are
CC useful for the isolation and detection of HPV nucleic acid sequences.

SQ Sequence 449 BP; 145 A; 78 C; 79 G; 137 T; 10 other;

Query Match 98.0%; Score 19.6; DB 18; Length 449;
Best Local Similarity 95.0%; Pred. No. 2.1;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 cactgaagtaactaagaag 20
Db 90 cactgaagtaactaagaag 109

RESULT 8

AAH87677
ID AAH87677 standard; CDNA; 606 BP.

AC AAH87677;

DT 25-SEP-2001 (first entry)

DE Peppermint plant oil gland expressed CDNA 33.

KM Peppermint; plant oil gland cell; terpenoid essential oil; resin;

OS genetic mapping; antisense suppression; recombinant expression; ss.

OS Mentha x piperita.

PN WO200153319-A1.

PF 19-JAN-2001; 2001WO-US02567.

PR 20-JAN-2000; 2000US-0177264.

PA (CROT/) CROTEAU R B.

PA (LANG/) LANGE B M.

PA (WILD/) WILDUNG M R.

PI Croteau RB, Lange BM, Wildung MR.

DR WPI: 2001-488706/53.

PT New nucleic acid molecules corresponding to mRNA molecules expressed in

PT peppermint oil glands for enhancing expression of plant oil gland cell

PT proteins

PS Claim 1; Page 89; 251pp; English.

CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
CC correspond to all or part of a mRNA molecule expressed in plant oil
CC gland cells, especially peppermint and plant oil glands that produce
CC terpenoid essential oils and resins. The nucleic acids are useful for
CC genetically mapping a plant genome for genes expressed in plant oil
CC gland cells and to suppress (for example by antisense suppression) or
CC enhance their expression (for example by genetically transforming a
CC plant cell with a replicable expression vector that expresses one or more
CC proteins naturally expressed in plant oil gland cells). The nucleic acids


```

XX (HOFF ) ROCHE MOLECULAR SYSTEMS INC.
PA
XX
XX Bauer HM, Gravitt PE, Greer CE, Impraim CC, Manos MM;
PI Resnick RM, Zhang TY;
XX
XX WPI: 1997-332084/30.
DR
XX
XX New oligo:nucleotide probes for human papilloma-virus - used for
PT detecting and typing HPV and for detecting previously unknown HPV
PT types and subtypes
XX
XX Disclosure: Columns 83-84: 94pp; English.
XX
XX The present sequence is a papillomavirus 251 (Pap251) specific
CC probe.
CC
XX
XX Sequence 25 BP; 10 A; 4 C; 5 G; 6 T; 0 other;
SQ

Query Match      80.0%; Score 16; DB 18; Length 25;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactag 16
DB 10 cactgaagtaactag 25

RESULT 14
AAT44888
ID AAT44888 standard; DNA: 46 BP.
XX
AC AAT44888;
XX
DT 05-FEB-1997 (first entry)
XX
DE HPV linked oligonucleotide probe TY2018.
XX
XX Probe; primer: PCR; polymerase chain reaction; amplification;
XX human papillomavirus; consensus; ss.
XX
XX Synthetic.
XX
XX US5527898-A.
XX
XX 18-JUN-1996.
XX
XX 09-SEP-1988; 88US-0243486.
XX
XX 24-SEP-1993; 93US-0126452.
XX 09-SEP-1988; 88US-0243486.
XX 10-MAR-1989; 89US-0322550.
XX 09-SEP-1989; 89MO-US03747.
XX 14-NOV-1990; 90US-0613142.
XX 20-APR-1993; 93US-0050743.
XX 07-JUN-1995; 95US-0474542.
XX
XX (HOFF ) HOFFMANN LA ROCHE INC.
XX
XX Bauer HM, Gravitt PE, Greer CE, Manos MM, Resnick RM;
PI Zhang TY;
XX
XX WPI: 1996-299903/30.
XX
XX Nucleic acid hybridisation probes - specific for selected human
PT papilloma virus types
XX
XX Example 7; Column 59-60; 96pp; English.
XX
XX The invention relates to new oligonucleotide probes and primers used
CC for the detection of human papillomaviruses (HPV) which are not genital
CC types 6, 11, 16, 18 or 33. The probes and primers AAT44608-T44693 are

```

```

CC esp. used to detect HPV types 26, 31, 31B, 35, 39, 40, 43, 45, 51-59 and
CC 68. The primers can be used to detect these HPV types in conjunction with
CC the consensus primers and typing probes AAT44733-T44906, which are based
CC on and amplify fragments of the L1, E6, E7 and E1 regions of the HPV
CC sequences. Detection of the amplification products is done with probes
CC derived from consensus sequences found in all characterised HPV
CC sequences.
CC Probes AAT44874-89 are probes used to detect amplified products from the
CC L1 region. Each probe contains two hybridising sequences specific to
CC each of 2 hypervariable regions of the L1 sequence. This results in a
CC doubling of probes available for detection by an immobilisation assay in
CC a microtitre plate. This probe is specific for HPV type 58.
XX
XX Sequence 46 BP; 20 A; 10 C; 9 G; 7 T; 0 other;
SQ

Query Match      80.0%; Score 16; DB 17; Length 46;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtaactag 16
DB 31 cactgaagtaactag 46

RESULT 15
AAT78109
ID AAT78109 standard; DNA: 46 BP.
XX
AC AAT78109;
XX
DT 07-OCT-1997 (first entry)
XX
DE Human papillomavirus 58 linked oligonucleotide probe TY2018.
XX
XX Human; papillomavirus 58; HPV58; linked oligonucleotide probe;
XX detection; ss.
XX
XX Synthetic.
XX
XX US5639871-A.
XX
XX 17-JUN-1997.
XX
XX 09-SEP-1988; 88US-0243486.
XX
XX 24-SEP-1993; 93US-0126452.
XX 09-SEP-1988; 88US-0243486.
XX 10-MAR-1989; 89US-0322550.
XX 09-SEP-1989; 89MO-US03747.
XX 14-NOV-1990; 90US-0613142.
XX 20-APR-1993; 93US-0050743.
XX 01-JUN-1995; 95US-0457648.
XX
XX (HOFF ) ROCHE MOLECULAR SYSTEMS INC.
XX
XX Bauer HM, Gravitt PE, Greer CE, Impraim CC, Manos MM;
PI Resnick RM, Zhang TY;
XX
XX WPI: 1997-332084/30.
XX
XX New oligo:nucleotide probes for human papilloma-virus - used for
PT detecting and typing HPV and for detecting previously unknown HPV
PT types and subtypes
XX
XX Example 7; Columns 153-154; 94pp; English.
XX
XX The present sequence is a human papillomavirus 58 (HPV58)
CC linked oligonucleotide probe.
CC
XX
XX Sequence 46 BP; 20 A; 10 C; 9 G; 7 T; 0 other;
SQ

```

```

Query Match: 80.0%; Score 16; DB 18; Length 46;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 16; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
Gy 1 cactgaatgaactag 16
    | | | | | | | | | |
Db 31 cactgaatgaactag 46

```

Search completed: April 3, 2002, 11:05:29
Job time: 5520 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 3, 2002, 10:31:07 ; Search time 143.87 seconds
(without alignments)
31.484 Million cell updates/sec

Title: US-09-885-799-318

Sequence: 1 cactgaagtaactaaggaag 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCRTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	22	1	US-08-050-743-72
2	20	100.0	22	1	US-08-474-542A-220
3	20	100.0	22	1	US-08-457-648-220
4	20	100.0	22	1	US-08-452-055-72
5	20	100.0	139	4	US-08-739-103A-33
6	16.4	82.0	176373	3	US-09-128-153-17
7	16	80.0	25	1	US-08-474-542A-59
8	16	80.0	25	1	US-08-457-648-59
9	16	80.0	46	1	US-08-474-542A-245
10	16	80.0	46	1	US-08-457-648-245
11	15.2	76.0	975	3	US-08-216-001-2
12	15.2	76.0	975	3	US-08-818-862-2
13	15.2	76.0	3018	3	US-08-942-572-1
14	14.8	74.0	4140	3	US-08-894-731-2
15	14.4	72.0	1424	4	US-09-613-182-13
16	14.4	72.0	4643	2	US-08-605-106-6
17	14.2	71.0	78	2	US-08-235-515A-3
18	14.2	71.0	141	2	US-08-235-515A-36
19	14.2	71.0	200	2	US-08-235-515A-56
20	14.2	71.0	373	2	US-08-791-849A-10
21	14.2	71.0	663	4	US-09-328-111-483
22	14.2	71.0	675	4	US-08-998-416-1127
23	14.2	71.0	707	4	US-08-998-416-1119
24	14.2	71.0	804	3	US-08-935-263-9
25	14.2	71.0	1386	2	US-08-408-095-22
26	14.2	71.0	1527	2	US-07-956-483-26
27	14.2	71.0	1527	4	US-08-472-240A-19

28	14.2	71.0	1676	1	US-08-484-105-21	Sequence 21, Appl
29	14.2	71.0	1676	1	US-08-484-106-21	Sequence 21, Appl
30	14.2	71.0	1892	4	US-09-423-340-3	Sequence 3, Appl
31	14.2	71.0	2454	3	US-08-890-615-1	Sequence 1, Appl
32	14.2	71.0	2475	3	US-09-045-185-1	Sequence 1, Appl
33	14.2	71.0	2475	3	US-09-045-185-3	Sequence 3, Appl
34	14.2	71.0	2999	2	US-09-014-969-8	Sequence 8, Appl
35	14.2	71.0	3157	6	5198347-3	Patent No. 5198347
36	14.2	71.0	3268	3	US-09-356-952-13	Sequence 13, Appl
37	14.2	71.0	5892	3	US-08-755-587-27	Sequence 27, Appl
38	14.2	71.0	6270	1	US-08-418-893D-25	Sequence 25, Appl
39	14.2	71.0	6371	3	US-08-836-325-13	Sequence 13, Appl
40	14.2	71.0	6404	3	US-08-836-325-14	Sequence 14, Appl
41	14.2	71.0	6790	1	US-08-418-893D-22	Sequence 22, Appl
42	14.2	71.0	7240	3	US-08-755-587-15	Sequence 15, Appl
43	14.2	71.0	9737	2	US-08-944-449-7	Sequence 7, Appl
44	14.2	71.0	11385	2	US-08-639-501-1	Sequence 1, Appl
45	14.2	71.0	11385	3	US-09-044-946-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-050-743-72
; Sequence 72, Application US/08050743
; Patent No. 5447839
; GENERAL INFORMATION:
; APPLICANT: Bauer, Heidi M.
; APPLICANT: Greer, Catherine E.
; APPLICANT: Manos, Michele
; APPLICANT: Resnick, Robert M.
; APPLICANT: Ting, Yi
; TITLE OF INVENTION: Detection of Human Papillomavirus by the
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,743
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sias, Stacey R.
; REGISTRATION NUMBER: 32,630
; REFERENCE/DOCKET NUMBER: 8793
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 814-2863
; TELEFAX: (510) 814-2977
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-050-743-72

Query Match 100.0%; Score 20; DB 1; Length 22;
Best local Similarity 100.0%; Pred. No. 0.35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 cactgaagtaactaaggaag 20
        |||||
Db       2 CACTGAAGTACTAAGGAAG 21
```

RESULT 2
US-08-474-542A-220
; Sequence 220, Application US/08474542A
; Patent No. 5527898

GENERAL INFORMATION:

APPLICANT: Bauer, Heidi M.
APPLICANT: Gravitt, Patti E.

APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.

TITLE OF INVENTION: Polymerase Chain Reaction
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.

STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110

```

1  SOFTWARE:  PatentIn Release #1.0, Version #1.25
2
3  CURRENT APPLICATION DATA:
4
5  APPLICATION NUMBER:  US/08/474,542A
6
7  FILING DATE:

```

Query Match	100.08;	Score 20;	DB 1;	Length 22
Best Local Similarity	100.08;	Pred. No. 0.35;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels	

RESULT 3
US-08-457-648-220
; Sequence 220, Application US/08457648
; Patent No. 5639871

GENERAL INFORMATION:

;
; APPLICANT: Bauer, Heidi M.
; APPLICANT: Gravitt, Patli E.

INVENTOR: HANOS, M. MICHAEL
 APPLICANT: Resnick, Robert M.
 TITLE OF INVENTION: Detection of Human Papillomavirus by the

TITLE OF INVENTION Polymerase Chain Reaction
 NUMBER OF SEQUENCES: 298
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hoffmann-La Roche Inc.
 STREET: 340 Kingsland Street
 CITY: Nutley

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy
;

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION

```

Query Match      100.0%; Score 20; DB 1; length 22;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 4
US-08-452-055-72
; Sequence 72, Application US/08452055

GENERAL INFORMATION

APPLICANT: Bauer, Heidi M.
APPLICANT: Greer, Catherine E

APPLICANT: Resnick, Robert M.

;	TITLE OF INVENTION:	Detection of Human Papillomavirus by the
;	TITLE OF INVENTION:	Polymerase Chain Reaction

```

; ZIP: 0/110
; COMPUTER READABLE FORM
;

```


ATTORNEY/AGENT INFORMATION:
NAME: Sias, Slacey R.
REGISTRATION NUMBER: 32,630
REFERENCE/DOCKET NUMBER: 9188
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2863
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-452-055-72

Query Match 100.0%; Score 20; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtactaagaag 20
|||||
DB 2 CACTGAAGTACTAAGAAG 21

RESULT 5

US-08-739-103A-33
Sequence 33, Application US/08739103A
Patent No. 6265154
GENERAL INFORMATION:
APPLICANT: P. Kroegeer
APPLICANT: K. Abiravaya
APPLICANT: J. Gorzowski
APPLICANT: R. Hoehnle
APPLICANT: J. Moore
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FOR DETECTING
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.0.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/739,103A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Paul D. Yasger
REGISTRATION NUMBER: 37,477
REFERENCE/DOCKET NUMBER: 5990,US-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-2341
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA (HPV type 58)
US-08-739-103A-33

Query Match 100.0%; Score 20; DB 4; Length 139;

Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cactgaagtactaagaag 20
|||||
DB 48 CACTGAAGTACTAAGAAG 67

RESULT 6

US-09-128-155-17/c
Sequence 17, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: fastseq for windows version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 82.0%; Score 16.4; DB 3; Length 176373;
Best Local Similarity 94.4%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 actgaagtactaagaag 19
|||||
DB 8025 ACTGAAGTACTAAGAAG 8008

RESULT 7

US-08-474-542A-59
Sequence 59, Application US/08474542A
Patent No. 5527898
GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravitt, Paul E.
APPLICANT: Greer, Catherine E.
APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSER: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,542A

FILED DATE: 435
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9234
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-542A-59

Query Match 80.0%; Score 16; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cactgaagtaactaag 16
|||||
DB 10 CACTGAAGTAAGTAAG 25

RESULT 8
US-08-457-648-59
Sequence 59, Application US/08457648
Patent No. 5639871
GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravit, Patti E.
APPLICANT: Greer, Catherine E.
APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,648
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-457-648-59

Query Match 80.0%; Score 16; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cactgaagtaactaag 16
|||||
DB 10 CACTGAAGTAAGTAAG 25

RESULT 9
US-08-474-542A-245
Sequence 245, Application US/08474542A
Patent No. 5527898
GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravit, Patti E.
APPLICANT: Greer, Catherine E.
APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,542A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9234
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 245:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-542A-245

Query Match 80.0%; Score 16; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cactgaagtaactaag 16
|||||
DB 31 CACTGAAGTAAGTAAG 46

RESULT 10
US-08-457-648-245
Sequence 245, Application US/08457648
Patent No. 5639871
GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.

APPLICANT: Gravitt, Patti E.
APPLICANT: Greer, Catherine E.
APPLICANT: Ingram, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the
NUMBER OF SEQUENCES: 298
TITLE OF INVENTION: Polymerase Chain Reaction
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,648
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pelety, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ. ID NO.: 245:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-457-648-245

Query Match 80.0%; Score 16; DB 1; Length 46;
Best local Similarity 100.0%; Pred. No. 27;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cactgaagtaactaag 16
DB 31 CACTGAAGTACTAAG 46
RESULT 11
US-09-216-001-2/c
Sequence 2, Application US/09216001
Patent No. 6004792
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Purvi
APPLICANT: Corley, Lynn E.
TITLE OF INVENTION: NEW HUMAN LYSOPHOSPHOLIPASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/216,001
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/878,862
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0329 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 975 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYMNOT03
CLONE: 2554166
US-09-216-001-2

Query Match 76.0%; Score 15.2; DB 3; Length 975;
Best local Similarity 85.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 cactgaagtaactaaggaag 20
DB 528 CACTGAAGTACTAAGGAG 509
RESULT 12
US-08-878-862-2/c
Sequence 2, Application US/08878862
Patent No. 613544
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Purvi
APPLICANT: Corley, Lynn E.
TITLE OF INVENTION: NEW HUMAN LYSOPHOSPHOLIPASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,862
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0329 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

TELEX:
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 975 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? IMMEDIATE SOURCE:
? LIBRARY: THYNOT03
? CLONE: 2554166
US-08-878-862-2

Query Match 76.0%; Score 15.2; DB 3; Length 975;
Best Local Similarity 85.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 cactgaagtaactaagaaq 20
||||| ||| |||||
Db 528 CACTGAATATTCAGGAG 509

RESULT 13
US-08-942-572-1/C
? Sequence 1, Application US/08942572
? Patent No. 6140488
? GENERAL INFORMATION:
? APPLICANT: Avruich, Joseph
? APPLICANT: Zhang, Xian-Feng
? TITLE OF INVENTION: RAS-BINDING PROTEIN
? NUMBER OF SEQUENCES: 2
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson, P.C.
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: MA
? COUNTRY: US
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FastSeq for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/942,572
? FILING DATE: 01-OCT-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Fraser, Janis K.
? REGISTRATION NUMBER: 34,819
? REFERENCE/DOCKET NUMBER: 08472/737001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617/542-5070
? TELEFAX: 617/542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3018 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 31...1269
US-08-942-572-1

Query Match 76.0%; Score 15.2; DB 3; Length 3018;
Best Local Similarity 85.0%; Pred. No. 77;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 cactgaagtaactaagaaq 20
||||| ||| |||||
Db 2303 CACTGAAGTACACAGGAG 2284

RESULT 14
US-08-894-731-2/C
? Sequence 2, Application US/08894731
? Patent No. 6084089
? GENERAL INFORMATION:
? APPLICANT: MINE, Toshiki
? APPLICANT: OHYAMA, AKIO
? APPLICANT: HAYASHI, Toru
? APPLICANT: KASAKA, Keisuke
? TITLE OF INVENTION: COLD-INDUCIBLE PROMOTER SEQUENCE
? FILE REFERENCE: 760-234P
? CURRENT APPLICATION NUMBER: US/08/894,731
? CURRENT FILING DATE: 1997-10-27
? NUMBER OF SEQ ID NOS: 8
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 2
? LENGTH: 4140
? TYPE: DNA
? ORGANISM: Solanum tuberosum
US-08-894-731-2

Query Match 74.0%; Score 14.8; DB 3; Length 4140;
Best Local Similarity 88.9%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 actgaagtaactaagaa 19
| ||||| |||||
Db 770 ATTGAAGTACTCAGGAA 753

RESULT 15
US-09-613-182-13/C
? Sequence 13, Application US/09613182
? Patent No. 6294653
? GENERAL INFORMATION:
? APPLICANT: Mayfield, Stephen
? TITLE OF INVENTION: RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR
? FILE REFERENCE: SCR2177S
? CURRENT APPLICATION NUMBER: US/09/613,182
? CURRENT FILING DATE: 2000-07-10
? PRIOR APPLICATION NUMBER: 09/341,550
? PRIOR FILING DATE: 1999-07-13
? PRIOR APPLICATION NUMBER: PCT/US98/00840
? PRIOR FILING DATE: 1998-01-16
? PRIOR APPLICATION NUMBER: 60/035,955
? PRIOR FILING DATE: 1997-01-17
? PRIOR APPLICATION NUMBER: 60/069,400
? PRIOR FILING DATE: 1997-12-12
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 13
? LENGTH: 1424
? TYPE: DNA
? ORGANISM: Chlamydomonas reinhardtii
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (252)..(1310)
? NAME/KEY: misc_feature
? LOCATION: (279)
? OTHER INFORMATION: Codon also can encode Ser
? NAME/KEY: misc_feature
? LOCATION: (282)
? OTHER INFORMATION: Codon also can encode Glu
? NAME/KEY: misc_feature

LOCATION: (294)
 OTHER INFORMATION: Codon also can encode Gly
 NAME/KEY: misc_feature
 LOCATION: (306)
 OTHER INFORMATION: Codon also can encode Asn
 NAME/KEY: misc_feature
 LOCATION: (357)
 OTHER INFORMATION: Codon also can encode Leu
 NAME/KEY: misc_feature
 LOCATION: (369)
 OTHER INFORMATION: Codon also can encode Thr
 NAME/KEY: misc_feature
 LOCATION: (486)
 OTHER INFORMATION: Codon also can encode Ser
 NAME/KEY: misc_feature
 LOCATION: (495)
 OTHER INFORMATION: Codon also can encode Ile
 NAME/KEY: misc_feature
 LOCATION: (510)
 OTHER INFORMATION: Codon also can encode Ala
 NAME/KEY: misc_feature
 LOCATION: (555)
 OTHER INFORMATION: Codon also can encode Val
 NAME/KEY: misc_feature
 LOCATION: (588)
 OTHER INFORMATION: Codon also can encode Gly
 NAME/KEY: misc_feature
 LOCATION: (600)
 OTHER INFORMATION: Codon also can encode Leu
 NAME/KEY: misc_feature
 LOCATION: (621)
 OTHER INFORMATION: Codon also can encode Ala
 NAME/KEY: misc_feature
 LOCATION: (714)
 OTHER INFORMATION: Codon also can encode Thr
 NAME/KEY: misc_feature
 LOCATION: (729)
 OTHER INFORMATION: Codon also can encode Ile
 NAME/KEY: misc_feature
 LOCATION: (1092)
 OTHER INFORMATION: Codon also can encode Val
 NAME/KEY: misc_feature
 LOCATION: (1191)
 OTHER INFORMATION: Codon also can encode Ile
 NAME/KEY: misc_feature
 LOCATION: (1284)
 OTHER INFORMATION: Codon also can encode Ala
 NAME/KEY: misc_feature
 LOCATION: (1287)
 OTHER INFORMATION: Codon also can encode Ile
 NAME/KEY: misc_feature
 LOCATION: (1290)
 OTHER INFORMATION: Codon also can encode Glu
 NAME/KEY: misc_feature
 LOCATION: (1293)
 OTHER INFORMATION: Codon also can encode Ala
 NAME/KEY: misc_feature
 LOCATION: (1296)
 OTHER INFORMATION: Codon also can encode Pro
 NAME/KEY: misc_feature
 LOCATION: (1302)
 OTHER INFORMATION: Codon also can encode Thr
 NAME/KEY: misc_feature
 LOCATION: (1308)
 OTHER INFORMATION: Codon also can encode Gly
 US-09-613-182-13

Query Match 72.0%; Score 14.4; DB 4; Length 1424;
 Best Local Similarity 93.8%; Pred. No. 1.8e+02;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 tgaagtaactaagaa 19

Db 914 TGAAGTAAGTAAAGAA 899

Search completed: April 3, 2002, 10:31:11
 Job time: 3462 sec

Thu Apr 4 09:39:25 2002

us-09-885-799-318.rni

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 10:28:33 ; Search time 3285.41 Seconds

(without alignments)
65.415 Million cell updates/sec

Title: US-09-885-799-318

Sequence: 1 cactgaagtaactaagaag 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: em_estfun:
2: em_estnum:
3: em_estin:
4: em_estom:
5: em_estpl:
6: em_estba:
7: em_estro:
8: em_estrov:
9: em_hic:
10: gb_est1:
11: gb_est2:
12: gb_hic:
13: gb_gss:
14: em_gss_fun:
15: em_gss_hum:
16: em_gss_inv:
17: em_gss_pln:
18: em_gss_pro:
19: em_gss_rtd:
20: em_gss_vrt:
21: em_gss_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	92.0	547	10	AM624074
2	18	90.0	585	10	BE304278
3	17.4	87.0	357	10	BE231490
4	17.4	87.0	580	13	AQ348400
5	17.4	87.0	613	13	AZ402774
6	17.4	87.0	614	13	AQ593333
7	17.4	87.0	924	13	AZ687649
8	17	85.0	768	11	BC221969
9	16.8	84.0	277	10	AA347960
10	16.8	84.0	307	10	AA324900
11	16.8	84.0	310	10	AA308882
12	16.8	84.0	324	11	T16859

C 13	16.8	84.0	326	11	H43832	H43832	yp21e09.r1
C 14	16.8	84.0	334	10	AA908640	AA908640	cl04a02.s
C 15	16.8	84.0	334	11	BI227287	BI227287	602948748
C 16	16.8	84.0	358	11	T58103	T58103	yb26c03.r1
C 17	16.8	84.0	365	10	AA241722	AA241722	MBL2SJF8
C 18	16.8	84.0	369	11	BG008631	BG008631	PM4-GN030
C 19	16.8	84.0	394	11	BG677300	BG677300	602624012
C 20	16.8	84.0	397	11	R50310	R50310	yj61009.r1
C 21	16.8	84.0	401	11	BF197720	BF197720	7n61905.x
C 22	16.8	84.0	449	11	H43503	H43503	yp98b06.r1
C 23	16.8	84.0	449	10	AA48564	AA48564	zx09a09.r
C 24	16.8	84.0	458	10	AA448610	AA448610	zx09e11.r
C 25	16.8	84.0	470	10	AA054781	AA054781	2K68d02.r
C 26	16.8	84.0	470	10	AA150511	AA150511	2108a01.r
C 27	16.8	84.0	483	10	AA808747	AA808747	oe56h06.s
C 28	16.8	84.0	494	10	AM649401	AM649401	EST327855
C 29	16.8	84.0	507	10	AM649401	AM649401	EST327855
C 30	16.8	84.0	507	10	AM649401	AM649401	EST327855
C 31	16.8	84.0	510	11	N98715	N98715	yp65f07.r1
C 32	16.8	84.0	514	10	AM649148	AM649148	EST327602
C 33	16.8	84.0	541	10	AM221164	AM221164	EST297633
C 34	16.8	84.0	562	10	AM617871	AM617871	EST324282
C 35	16.8	84.0	573	10	BC134037	BC134037	EST466929
C 36	16.8	84.0	580	13	BH083071	BH083071	RPCL-24-4
C 37	16.8	84.0	582	10	AA933378	AA933378	SMBM135BH
C 38	16.8	84.0	592	10	A1781400	A1781400	EST262267
C 39	16.8	84.0	604	10	BC354705	BC354705	MBACAP4B
C 40	16.8	84.0	606	10	AM624961	AM624961	EST313790
C 41	16.8	84.0	618	10	AL037896	AL037896	ML234 pep
C 42	16.8	84.0	622	10	AW038473	AW038473	EST280156
C 43	16.8	84.0	622	10	A1442791	A1442791	sa26e12.x
C 44	16.8	84.0	623	10	A1775250	A1775250	EST256350
C 45	16.8	84.0	623	10	AM038046	AM038046	EST279703

ALIGNMENTS

RESULT 1	AM624074	547 bp	MRNA	EST	18-MAY-2001
LOCUS	AM624074				
DEFINITION	EST322019 tomato flower buds 3-8 mm, Cornell University				
ACCESSION	Lycopersicon esculentum cDNA clone CT0B14D21 5', mRNA sequence.				
VERSION	AM624074.1	GI:7337101			
KEYWORDS	EST.				
SOURCE	tomato.				
ORGANISM	Lycopersicon esculentum				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;				
AUTHORS	Lycopersicon				
	1 (bases 1 to 547)				
	van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Romling, C.M., Niernann, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J., and Tanksley, S.D.				
	Generation of ESTs from tomato flower tissue, 3-8 mm buds				
	Unpublished (1999)				
COMMENT	Contact: CUGI				
	Clemson University Genomics Institute				
	100 Jordan Hall, Clemson, SC 29634, USA				
	Email: http://www.genome.clemson.edu/orders/index.html				
	5 prime sequence.				
FEATURES	Location/Qualifiers				
SOURCE	1..547				
	/organism="Lycopersicon esculentum"				
	/cultivar="TA496"				
	/db_xref="taxon:4081"				
	/clone="CT0B14D21"				
	/clone_id="tomato flower buds 3-8 mm, Cornell University"				
	/tissue_type="flower"				
	/dev_stage="3-8mm buds"				

/note="Vector: pBluescript SK(-): Site.1: EcoRI; Site.2: XhoI; supplier: Tankley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, T456). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT 162 a 95 c 126 g 164 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 547;
Best Local Similarity 95.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaa 20
|||||
Db 522 CACTGAAGTAAGTAAGAGC 541

RESULT 2

BE304278 585 bp mRNA EST 13-JUL-2000
LOCUS 601086085P1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3500253 5',
DEFINITION RNA sequence.
ACCESSION BE304278
VERSION BE304278.1 GI:9174558
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 585)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: L2A06358 row: 1 column: 22
High quality sequence stop: 582.

FEATURES
source
Location/Qualifiers
1..585
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="3500253"
/clone_id="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH."

BASE COUNT 176 a 116 c 152 g 141 t

ORIGIN

Query Match 90.0%; Score 18; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 actgaagtaactaagaa 19
|||||
Db 437 ACTGAAGTAAGTAAGCA 454

RESULT 3
BE231490/c
LOCUS BE231490
DEFINITION AP 48-r4 Pig anterior pituitary cDNA Sus scrofa cDNA, mRNA
ACCESSION BE231490
VERSION BE231490
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa

REFERENCE 1 (bases 1 to 357)
AUTHORS Bertani, G.R., Johnson, R.K. and Pomp, D.
TITLE Gene expression studies in anterior pituitary of pigs selected for fertility
JOURNAL Unpublished (2000)
COMMENT Contact: Bertani GR
Animal Science Molecular Genetics Lab
University of Nebraska
Animal Science A218c, Lincoln, NE 68583-0908, USA
Tel: 402/4724515
Fax: 402/4726362
Email: gbertani@bigred.unl.edu
novel
PCR Primers
FORWARD: ACAATTTCACACAGATCCATGATC (ARBITRARY4)
BACKWARD: ACCACTCCTATAGGCGTTTCTTTTTCGC (ANCHOR2)
Seq primer: M13forward -21
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1..357
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="Pig anterior pituitary cDNA"
/sex="female"
/tissue_type="anterior pituitary"
/dev_stage="follicular phase"
/note="EST's isolated from differential display experiment"

BASE COUNT 132 a 55 c 52 g 118 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 357;
Best Local Similarity 94.7%; Pred. No. 6.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cactgaagtaactaagaa 19
|||||
Db 56 CACTGAAGTAAGTAAGCA 38

RESULT 4
A0348400 580 bp DNA GSS 07-MAY-1999
LOCUS R03111-125D13.TV R03111 Homo sapiens genomic clone R03111-125D13,
DEFINITION DNA sequence.
ACCESSION A0348400
VERSION A0348400.1 GI:4173296
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 580)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
TITLE Use of BAC End Sequences from Library R03111 for Sequence-Ready Map Building (1997)
JOURNAL Unpublished
COMMENT other_GSSs: R03111-125D13.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetj@igrr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (<http://inforesgen.com>). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: 17
Class: BAC ends.

FEATURES

Source Location/Qualifiers

1.580
/organism="Homo sapiens"
/db_xref="GDB:7547700"
/db_xref="taxon:9606"
/clone="RPCI-11-125D13"
/clone_1b="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPC11 Human Male BAC Library"
BASE COUNT 175 a 112 c 103 g 187 t 3 others
ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 580;
Best Local Similarity 94.7%; Pred. No. 6.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 actgaagtaactaagaag 20
|||||
Db 279 ACTGACGTACTTAAGAG 297

RESULT 5

AZ402774

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 613)

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0170 Row: B Column: 18

Seq primer: CGTGTAAACGACGCGCACT

Class: plasmid ends

High quality sequence stop: 613.

Location/Qualifiers

1.613

Source

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0170B18"
/clone_1b="Mouse 10kb plasmid UUC1M library"
/sex="Male"

/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD29v; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (gil473211419b1AR129072.1), a copy number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 196 a 173 c 134 g 109 t 1 others
ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 613;
Best Local Similarity 94.7%; Pred. No. 6.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 actgaagtaactaagaag 20
|||||
Db 63 ACTGACGTACTTAAGAG 81

RESULT 6

A0593333

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 614)

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

Contact: Mahairas GG, Wallace JC, Hood L

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (<http://inforesgen.com>). BAC end web Server: <http://www.htsc.washington.edu>

Plate: 1030 Row: H Column: 24

Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 614.
 Location/Qualifiers
 1. 614

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="plate=1030 COL=24 Row=H"
 /sex="male"
 /note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBACE3.6 vector at EcoRI sites"

BASE COUNT 164 a 128 c 119 g 186 t 17 others
 ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 614;
 Best Local Similarity 94.7%; Pred. No. 6.4e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 actgaagtaactaagaag 20
 |||||
 Db 335 ACTGAGTAATTAAGCAAG 353

RESULT 7
 A2687649/c 924 bp DNA GSS 14-DEC-2000
 LOCUS
 DEFINITION EMPVC85TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
 ACCESSION A2687649
 VERSION A2687649.1 GI:11824795
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica.
 ORGANISM Entamoeba histolytica.
 Eukaryota; Entamoebidae; Entamoeba.
 REFERENCE 1 (bases 1 to 924)
 AUTHORS Loftus, B., Van Aken, S., and Fraser, C.
 TITLE Determination of clone end sequences from Entamoeba histolytica
 JOURNAL Unpublished (2000)
 COMMENT Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: entactgr.org
 Clones are derived from the Entamoeba histolytica HML:IMSS sheared
 DNA library
 Seq primer: M13-Reverse
 Class: Shotgun
 High quality sequence start: 27
 High quality sequence stop: 773.
 Location/Qualifiers
 1. 924

FEATURES
 SOURCE

/organism="Entamoeba histolytica"
 /strain="HML:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: PHOS1; Site 1: Bst I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450. The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + i method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome
 Sequencing: A Practical Approach, eds. M. Vaubin and B.
 Borell, Oxford University Press, 1993).

BASE COUNT 379 a 100 c 158 g 287 t
 ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 924;
 Best Local Similarity 94.7%; Pred. No. 6.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 actgaagtaactaagaag 20
 |||||
 Db 857 ACTGAGTAATTAAGCAAG 839

RESULT 8
 BG221969 768 bp mRNA EST 21-APR-2001
 LOCUS
 DEFINITION R5141787 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG221969
 VERSION BG221969.1 GI:13747990
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 768)
 AUTHORS Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
 Lerner, S., Dahl, T., Thornton, M., Ramchandran, R., Harrington, J.,
 Veloso, N., Hess, J., McElligott, K., Clark, S., Mays, R., Smith, E.,
 and Ducar, M.
 TITLE Creation of Genome-wide Protein Expression Libraries using Random
 Activation of Gene Expression
 JOURNAL Nat. Biotechnol. 19 (5), 440 (2001) In press
 COMMENT Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@atersys.com
 High quality sequence stop: 519.
 Location/Qualifiers
 1. 768

FEATURES
 SOURCE

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Athersys RAGE Library"
 /cell_line="HT1080"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in HT1080 under normal circumstances."
 BASE COUNT 269 a 108 c 149 g 242 t
 ORIGIN

Query Match 85.0%; Score 17; DB 11; Length 768;
 Best Local Similarity 100.0%; Pred. No. 9.7e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 actgaagtaactaaga 18
 |||||
 Db 403 ACTGAGTAATTAAGCA 419

RESULT 9
 AA347960/c 277 bp mRNA EST 21-APR-1997
 LOCUS
 DEFINITION EST54368 Fetal heart II Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA347960

VERSION AA347960.1 GI:2000217
 EST
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 277)

AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulner,R.A., Bult

,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., Bult
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Barle-Hughes,J., Fine,L.D., Fitzgerald
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodex,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudex,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A., He,W.M.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280

COMMENT Other ESTs: TH0171208

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423

Email: arkerlavet@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES

Source

1..277

Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="ATCC (inhost):149038"
 /db_xref="taxon:9606"
 /clone_lib="Fetal heart 11"
 /dev_stage="fetus"
 /note="Organ: heart; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"
 BASE COUNT 64 a 59 c 67 g 85 t 2 others
 ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 277;
 Best Local Similarity 90.0%; Pred. No. 1.2e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 cactgaagtaactgaagaag 20
 Db 89 CACTGAAGTAAGTCTGGAAG 70

RESULT 10
 AA324900/c 307 bp mRNA EST 20-APR-1997
 LOCUS EST17802 Cerebellum 11 Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA324900
 VERSION AA324900.1 GI:1977144
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 307)

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulner,R.A., Bult
 ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Barle-Hughes,J., Fine,L.D., Fitzgerald
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodex,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudex,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A., He,W.M.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280

COMMENT

Other ESTs: TH0171208

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423

Email: arkerlavet@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES

Source

1..307

Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="ATCC (inhost):125993"
 /db_xref="taxon:9606"
 /clone_lib="Cerebellum 11"
 /tissue_type="Cerebellum"
 /dev_stage="adult"
 /note="Organ: brain; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"
 BASE COUNT 69 a 67 c 77 g 91 t 3 others
 ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 307;
 Best Local Similarity 90.0%; Pred. No. 1.2e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 cactgaagtaactgaagaag 20
 Db 31 CACTGAAGTAAGTCTGGAAG 12

RESULT 11
 AA308882/c 310 bp mRNA EST 18-APR-1997
 LOCUS EST179686 Cerebellum 11 Homo sapiens cDNA 5' end similar to EST
 ACCESSION AA308882
 VERSION AA308882.1 GI:1961220
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 310)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulner,R.A., Bult
 ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,

Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritschman, J.L., Geoghagen, N.S., Glodok, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S. Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.B., Scott, J.L., Saudak, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dinke, D., Feng, D.-F., Ferlie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL

MEDLINE

COMMENT

96026280

Other ESTs: THC171208

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@igr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

SOURCE

1. 310

/db_xref="taxon:9606"

/db_xref="taxon:9606"

/clone_lib="Cerebellum 11"

/tissue_type="Cerebellum"

/dev_stage="adult"

/note="Organ: brain; Vector: pBluescript SK-; Site_1:

EcotRI; Site_2: XhoI"

BASE COUNT

66 a 66 c 81 g 97 t

ORIGIN

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 10; Length 310;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

REFERENCE

AUTHORS

1 cactgaagtaactgaaggaag 20

|||||

Db 143 CACTGAAGTACTGGAAG 124

RESULT 12

LOCUS

TI6859 324 bp mRNA

DEFINITION

NT16944-5R Normalized infant brain, Bento Soares Homo sapiens cDNA

5' end, mRNA sequence.

TI6859

VERSION

TI6859.1 GI:519021

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 324)

Berry, R., Stevens, T.J., Walter, N.A.R., Wilcox, A.S., Rubano, T.,

Hopkins, J.A., Weber, J., Gold, R., Soares, M.B. and Sikeia, J.M.

Gene-based Sequence Tagged Sites (STS) as the basis for a human

gene map

Nature Genet. 10, 415-423 (1995)

Contact: Sikeia JM

JOURNAL

MEDLINE

95400322

COMMENT

Contact: Sikeia JM

Department of Pharmacology
University of Colorado Health Sciences Center
Box C236, 4200 E. 9th Ave, Denver CO 80262-0236
Tel: 3032708637
Fax: 3032707097
Email: nikki@lady.uchsc.edu
Seq primer: M13 Reverse.

FEATURES

SOURCE

1. 324

/db_xref="taxon:9606"

/clone_lib="Normalized infant brain, Bento Soares"

/lab_host="E. coli DH5-alpha"

/note="Vector: BA, M13-derived; Site_1: HindIII; Site_2:

NotI; The normalized infant brain library, constructed by

Bento Soares, Columbia University, was oligo-(dT) primed

and directionally cloned into an M13-derived plasmid using

total brain mRNA from a 72-day old human female afflicted

with spinal muscular atrophy. The library was normalized

as described elsewhere."

BASE COUNT

64 a 76 c 78 g 102 t 4 others

ORIGIN

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 11; Length 324;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

REFERENCE

AUTHORS

1 cactgaagtaactgaaggaag 20

|||||

Db 275 CACTGAAGTACTGGAAG 256

RESULT 13

LOCUS

H43832 326 bp mRNA

DEFINITION

YP21609.r1 Soares breast 3NBHst Homo sapiens cDNA clone

IMAGE:188104 5' similar to contains Alu repetitive element; mRNA

sequence.

H43832

VERSION

H43832.1 GI:919684

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 326)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,

R., Williamson, A., Wohlmann, P. and Wilson, R.

The WashU-Merck EST project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

High quality sequence strops: 318

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: M13RP1

High quality sequence strop: 318.

Location/Qualifiers

1. 326

/db_xref="taxon:9606"

/clone_lib="Soares breast 3NBHst"

/sex="Female"

```

1. 334
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1522442"
/clone.lib="NCI_CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pMT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
neuroendocrine lung carcinoid, and was then primed with a
Not I - oligo(dT) primer. Double-stranded cDNA was ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pMT73 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo."
76 a 75 c 81 g 102 +

```

search completed: April 3, 2002, 10:28:35
 job time: 3306 sec

Thu Apr 4 09:39:27 2002

us-09-885-799-318.rst

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 09:33:29 ; Search time 1711.61 Seconds
(without alignments)
192.768 Million cell updates/sec

Title: US-09-885-799-317

Perfect score: 20

Sequence: 1 atgcacgaaglaactaag 20

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_com:*
20: em_ov:*
21: em_or:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hgo_hum:*
31: em_hgo_inv:*
32: em_hgo_hum:*
33: em_hgo_hum:*
34: em_hgo_inv:*
35: em_hgo_inv:*
36: em_hgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	30	6	A46153	A46153 Sequence 4
2	20	100.0	449	14	HP045924	U45924 Human pap
3	20	100.0	449	14	HP045925	U45925 Human pap
4	20	100.0	449	14	HP045926	U45926 Human pap
5	20	100.0	449	14	HP045927	U45927 Human pap
6	20	100.0	449	14	HP045928	U45928 Human pap
7	20	100.0	7824	14	PPH58	D90400 Human pap
8	19	95.0	25	6	122571	122571 Sequence 5
9	19	95.0	25	6	147396	147396 Sequence 5
10	19	95.0	46	6	122757	122757 Sequence 2
11	19	95.0	46	6	147582	147582 Sequence 2
12	18.4	92.0	449	14	HP045929	U45929 Human pap
13	18	90.0	22	6	114275	114275 Sequence 7
14	18	90.0	22	6	122732	122732 Sequence 2
15	18	90.0	22	6	147557	147557 Sequence 2
16	18	90.0	22	6	AF221510S8	AF221517 Sus scro
17	17.4	87.0	87166	2	AC014242	AC014242 Drosophi
18	17.4	87.0	130628	2	AC010992	AC010992 Drosophi
19	17.4	87.0	167276	9	AC025262	AC025262 Homo sap
20	17.4	87.0	173126	2	AC020623	AC020623 Homo sap
21	17.4	87.0	206525	2	AC022510	AC022510 Homo sap
22	17.4	87.0	323461	3	AE003491	AE003491 Drosophili
23	17	85.0	172203	9	AC005863	AC005863 Homo sapi
24	16.8	84.0	393	3	AB028276	AB028276 Gerris la
25	16.8	84.0	393	3	AB028277	AB028277 Gerris la
26	16.8	84.0	728	3	AB051395	AB051395 Gerris la
27	16.8	84.0	3209	9	BC001404	BC001404 Homo sapien
28	16.8	84.0	6640	9	D63997	D63997 Homo sapien
29	16.8	84.0	109770	2	AC073763-3	Continuation (4 of
30	16.8	84.0	113249	2	AC068805	AC068805 Mus muscu
31	16.8	84.0	120039	9	AL357124	AL357124 Human DNA
32	16.8	84.0	121682	2	AL139127	AL139127 Homo sapi
33	16.8	84.0	125661	9	AL138773	AL138773 Human DNA
34	16.8	84.0	143285	2	AC015996	AC015996 Homo sapi
35	16.8	84.0	152334	2	AC055864	AC055864 Homo sapi
36	16.8	84.0	152537	9	AC055731	AC055731 Homo sapi
37	16.8	84.0	156288	2	AC021069	AC021069 Homo sapi
38	16.8	84.0	160252	2	AL390250	AL390250 Homo sapi
39	16.8	84.0	160946	2	AC027671	AC027671 Homo sapi
40	16.8	84.0	161613	2	AC023947	AC023947 Homo sapi
41	16.8	84.0	164736	2	AC019344	AC019344 Homo sapi
42	16.8	84.0	167336	2	AC021350	AC021350 Homo sapi
43	16.8	84.0	167979	2	AL354770	AL354770 Homo sapi
44	16.8	84.0	172050	2	AL603784	AL603784 Homo sapi
45	16.8	84.0	176033	2	AL161794	AL161794 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS A46153 30 bp DNA
DEFINITION Sequence 48 from Patent WO9522626.
ACCESSION A46153
VERSION A46153.1 GI:2300401
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 30)
AUTHORS Meijer,C.J., Van,D.B., Walboomers,J.M. and Snijders,P.J.
TITLE HUMAN PAPILLOMA VIRUS DETECTION IN A NUCLEIC ACID AMPLIFICATION
JOURNAL PROCESS USING GENERAL PRIMERS
COMMENT Patent: WO 9522626-A 48 24-AUG-1995;
STICHTING RESEARCHFONDS PATROL (NL);
FEATURES Other publication AU 1672295 950904.
Location/Qualifiers
1..30
Source
/organism="unidentified"

PAT 07-MAR-1997

BASE COUNT 12 a 4 c 7 g 7 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgactgaagtaactaag 20
|||||
Db 4 ATGCACGAGTAAGTAAG 23

RESULT 2

HP045924 449 bp DNA VRL 20-NOV-1996
LOCUS
DEFINITION Human papillomavirus type 58, isolate IS068, major capsid protein
ACCESSION U45924
VERSION U45924.1 GI:1209174
KEYWORDS
SOURCE Human papillomavirus type 58.
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.

REFERENCE

AUTHORS Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.
TITLE Intratype variation in 12 human papillomavirus types: a worldwide
perspective
J. Virol. 70 (5), 3127-3136 (1996)

JOURNAL

MEDLINE 96186743
REFERENCE 2 (bases 1 to 449)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, HPV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA

COMMENT

IS068 was originally isolated from a cervical cancer specimen from
Brazil, in connection with the study published in Bosch, et al.,
J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this isolate has
been sequenced over the My09/11 consensus primer region, beginning
6588 bp downstream of the origin for HPV58. Due to primer
degeneracy, sequence information over the primer regions is
uncertain.

FEATURES

source Location/Qualifiers

1..449
/organism="Human papillomavirus type 58"
/isolate="IS068"
/db_xref="taxon:10598"

primer_bind

1..20
/gene="L1"
/note="My11 primer"
<1..>449

CDS

/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AAB19159.1"
/db_xref="GI:1209175"
/translation="AQGHNGICMGNOLEFVYVDTRSTNMLTCEVYKRGYKNDNF
KEYRVHVEEDLOFVFCOLCKITITAEIKYIHTMDSNILEDWQGLTPPSASLDDTY
RFVTSQAITCCKTAPKREKEDPLNKITFEVNLKEFSADLDQFLGR"
1..449
/gene="L1"
complement(429..449)
/note="My09 primer"

gene

primer_bind
BASE COUNT 150 a 79 c 83 g 137 t
ORIGIN

Query Match

100.0%; Score 20; DB 14; Length 449;

Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgactgaagtaactaag 20
|||||
Db 87 ATGCACGAGTAAGTAAG 106

RESULT 3

HP045925 449 bp DNA VRL 20-NOV-1996
LOCUS
DEFINITION Human papillomavirus type 58, isolate IS573, major capsid protein
ACCESSION U45925
VERSION U45925.1 GI:1209176
KEYWORDS
SOURCE Human papillomavirus type 58.
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.

REFERENCE

AUTHORS Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.
TITLE Intratype variation in 12 human papillomavirus types: a worldwide
perspective
J. Virol. 70 (5), 3127-3136 (1996)

JOURNAL

MEDLINE 96186743
REFERENCE 2 (bases 1 to 449)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, HPV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA

COMMENT

IS573 was originally isolated from a cervical cancer specimen from
Paraguay, in connection with the study published in Bosch, et al.,
J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this isolate
has been sequenced over the My09/11 consensus primer region,
beginning 6588 bp downstream of the origin for HPV58. Due to primer
degeneracy, sequence information over the primer regions is
uncertain.

FEATURES

source Location/Qualifiers

1..449
/organism="Human papillomavirus type 58"
/isolate="IS573"
/db_xref="taxon:10598"

primer_bind

1..20
/gene="L1"
/note="My11 primer"
<1..>449

CDS

/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AAB19160.1"
/db_xref="GI:1209177"
/translation="AQGHNGICMGNOLEFVYVDTRSTNMLTCEVYKRGYKNDNF
KEYRVHVEEDLOFVFCOLCKITITAEIKYIHTMDSNILEDWQGLTPPSASLDDTY
RFVTSQAITCCKTAPKREKEDPLNKITFEVNLKEFSADLDQFLGR"
1..449
/gene="L1"
complement(429..449)
/note="My09 primer"

primer_bind

BASE COUNT 150 a 78 c 84 g 137 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgactgaagtaactaag 20
|||||
Db 87 ATGCACGAGTAAGTAAG 106

RESULT 4
HP045926 449 bp DNA VRL 20-NOV-1996
LOCUS
DEFINITION Human papillomavirus type 58, isolate IS1021, major capsid protein
L1 (L1) gene, partial cds.
ACCESSION U45926
VERSION U45926.1 GI:1209178
KEYWORDS
SOURCE Human papillomavirus type 58.
ORGANISM Human papillomavirus type 58
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE 1 (bases 1 to 449)
AUTHORS Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.
TITLE Intratype variation in 12 human papillomavirus types: a worldwide
perspective
JOURNAL J. Virol. 70 (5), 3127-3136 (1996)
MEDLINE 96186743
REFERENCE 2 (bases 1 to 449)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, Hpv Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA
COMMENT IS1021 was originally isolated from a cervical cancer specimen from
al., J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this
isolate has been sequenced over the My09/11 consensus primer
region, beginning 6588 bp downstream of the origin for HPV58. Due
to primer degeneracy, sequence information over the primer regions
is uncertain.

FEATURES
source
1..449
/organism="Human Papillomavirus type 58"
/isolate="IS1021"
/db_xref="taxon:10598"
primer_bind
1..20
/gene="L1"
/note="My11 primer"
<1..>449
/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AAB19161.1"
/db_xref="GI:1209178"
/translation="AAGHNGICGNGQLFVTVDTTRSTNMTLCTEVTKEGTYKDNF
KEYVRHVEEDLDQFVQLCKITLTAEMVYIHTMSDILIEDMQLTPPESASLDY
RFVTSQAITCOKTAPPKKEDPLNKYTFWEVNLKEFSADLDQFPLGR"
1..449
/gene="L1"
complement(429..449)
primer_bind
/note="My09 primer"
BASE COUNT 148 a 78 c 86 g 137 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgcactgaactaag 20
|||||
Db 87 ATGCACCTGAAGTAAG 106

RESULT 5
HP045927 449 bp DNA VRL 20-NOV-1996
LOCUS
DEFINITION Human papillomavirus type 58, isolate IS131, major capsid protein

ACCESSION U45927
VERSION U45927.1 GI:1209180
KEYWORDS
SOURCE Human papillomavirus type 58.
ORGANISM Human papillomavirus type 58
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE 1 (bases 1 to 449)
AUTHORS Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.
TITLE Intratype variation in 12 human papillomavirus types: a worldwide
perspective
JOURNAL J. Virol. 70 (5), 3127-3136 (1996)
MEDLINE 96186743
REFERENCE 2 (bases 1 to 449)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, Hpv Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA
COMMENT IS131 was originally isolated from a cervical cancer specimen from
Bolivia, in connection with the study published in Bosch, et al.,
J. Natl. Cancer Inst. 87:796-802. A variant of HPV58, this isolate
has been sequenced over the My09/11 consensus primer region,
beginning 6588 bp downstream of the origin for HPV58. Due to primer
degeneracy, sequence information over the primer regions is
uncertain.

FEATURES
source
1..449
/organism="Human papillomavirus type 58"
/isolate="IS131"
/db_xref="taxon:10598"
primer_bind
1..20
/gene="L1"
/note="My11 primer"
<1..>449
/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AAB19162.1"
/db_xref="GI:1209181"
/translation="AAGHNGICGNGQLFVTVDTTRSTNMTLCTEVTKEGTYKDNF
KEYVRHVEEDLDQFVQLCKITLTAEMVYIHTMSDILIEDMQLTPPESASLDY
RFVTSQAITCOKTAPPKKEDPLNKYTFWEVNLKEFSADLDQFPLGR"
1..449
/gene="L1"
complement(429..449)
primer_bind
/note="My09 primer"
BASE COUNT 150 a 78 c 84 g 137 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 449;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgcactgaactaag 20
|||||
Db 87 ATGCACCTGAAGTAAG 106

RESULT 6
HP045928 449 bp DNA VRL 20-NOV-1996
LOCUS
DEFINITION Human papillomavirus type 58, isolate IS404, major capsid protein
L1 (L1) gene, partial cds.
ACCESSION U45928
VERSION U45928.1 GI:1209182
KEYWORDS
SOURCE Human papillomavirus type 58.
ORGANISM Human papillomavirus type 58


```

/note="open reading frame L2"
/codon_start=1
/protein_id="BAA31850.1"
/db_xref="GI:3337103"
/translation="MKRRSTRRRKASATDLYCTCKASCTCPDV1KPEVETIIADQI
LRYGSLGFPFGIGTGTGRTGTVPLGSPPEATPILOPIPPVATDTGPDLS
SIVLSIESSSPADAGAPAPISIPSPGDTTADTTPALINSSISGSIQVSTHN
PSTPEPSVLPPAPAPASGHI.IESSPVSHSENIMDTPIYISDSGNTSTPIPG
SKPVARLGLXSKNTQOVVDPALISPAKLVYTDPAFEGNPDITLOQSDISPA
PDDPDLIDIVALHRAPLISRGTVRSVWGKATLRSGKOIQAIVHYDLSPIQPV
QEVQOQOQOQELQSLNLSVSPYSINDLXDYADDDOTIHPOSLHSTSPATTS
NMSIPLNTGFDTPLVLSLEPCDIIASVTSMSSEPIISLPFMTIIVGDGMDLAPS
VFILRRRRKRPFFADVRAA"
5565..7139
/note="open reading frame L1"
/codon_start=1
/protein_id="BAA31851.1"
/db_xref="GI:3337104"
/translation="MVLICCTLALLFCVADVNVHIFLQMSVVRPSEATVYLPVPV
SKVSTDEVSTSTSYVAGSSRLAVGNPFYSIKSPNNKKVLYLPVYSGLOTRVRY
RLDPKRGFPPTSTSYNDQORLVMAVCGLEIGRQPLQVSGPYLTKFDDTETSN
RYPAQGSNRECLSMYKQTOULIGCKPPTGEMGKGVACNNMAATDCPELFLFN
SIIEDGMDVDTGFCGMDPGLQANKSDVPIDICNSTCKYPPYLMASEPYGDLFFFL
RROMFVRHFFENRAGKLEAVPDDLYIKSGWTAVIQSSAFEPYPPSGSIYTSSELEN
KPYMLQRAQHNGICQGNOLFTVVDTRSTNMTLCTEVRKSTYKNDMKRYVAV
EEYDLOFVQLCKITITAEIMTYITVMDTSNLEDDQFGLTPPSPASIODTRVTSQA
ITQCATPKEKEDPLNKYTFMEVNLKEFSADLDQPLCKKFLQSGIAKPRILARS
APTRAPSPKRRKKRYK"

```

BASE COUNT 2487 a 1388 c 1576 g 2373 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 7824;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 20
|||||
Db 6674 ATGCACGTGAAGTAAGTAAAG 6693

RESULT 8
LOCUS 122571 25 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 59 from patent US 5527898.
ACCESSION 122571
VERSION 122571.1 GI:1602925
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 25)
AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Manos,M.Michele, Resnick,R.M.
TITLE Detection of human papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5527898-A 59 18-JUN-1996;
FEATURES
source 1..25
BASE COUNT 10 a 4 c 5 g 6 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 19
|||||
Db 7 ATGCACGTGAAGTAAGTAAAG 25

RESULT 9

147396
LOCUS 147396 25 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 59 from patent US 5639871.
ACCESSION 147396
VERSION 147396.1 GI:2471361
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 25)
AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Impraim,C.C.,
Manos,M.Michele, Resnick,R.M. and Zhang,T.Y.
TITLE Detection of human papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5639871-A 59 17-JUN-1997;
FEATURES
source 1..25
BASE COUNT 10 a 4 c 5 g 6 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 19
|||||
Db 7 ATGCACGTGAAGTAAGTAAAG 25

RESULT 10
LOCUS 122757 46 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 245 from patent US 5527898.
ACCESSION 122757
VERSION 122757.1 GI:1603111
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 46)
AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Manos,M.Michele, Resnick,R.M.
TITLE Detection of human papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5527898-A 245 18-JUN-1996;
FEATURES
source 1..46
BASE COUNT 20 a 10 c 9 g 7 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 46;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 19
|||||
Db 28 ATGCACGTGAAGTAAGTAAAG 46

RESULT 11
LOCUS 147582 46 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 245 from patent US 5639871.
ACCESSION 147582
VERSION 147582.1 GI:2471547
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 46)
AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Impraim,C.C.,

TITLE
Journal
Patent: US 5639871-A-245 17-JUN-1997;
Location/Qualifiers

FEATURES

BASE COUNT 20 a 10 c 9 g 7 t
ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 46;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 19
|||||

HP045929 449 bp DNA VRL 20-NOV-1996
LOCUS Human papillomavirus type 58, isolate IS417, major capsid protein
DEFINITION L1 (L1) gene, partial cds.
ACCESSION U45929 GI:1209184
VERSION 1
KEYWORDS Human papillomavirus type 58.
SOURCE Human papillomavirus type 58.
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE 1 (bases 1 to 449)
AUTHORS Stewart,A.C., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
Peto,J. and Wheeler,C.M.,
TITLE Intratype variation in 12 human papillomavirus types: a worldwide
perspective
JOURNAL J. Virol. 70 (5), 3127-3136 (1996)
MEDLINE 96186743
REFERENCE 2 (bases 1 to 449)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew D. Farmer, Los Alamos National
Laboratory, HPV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87545, USA
COMMENT IS417 was originally isolated from a cervical cancer specimen from
Naili, in connection with the study published in Bosch, et al., J.
Natl. Cancer Inst. 87:795-802. A variant of HPV58, this isolate has
been sequenced over the My09/11 consensus primer region, beginning
6588 bp downstream of the origin for HPV58. Due to primer
degeneracy, sequence information over the primer regions is
uncertain.

FEATURES

1. 449 Location/Qualifiers
source /organism="Human papillomavirus type 58"
/isolate="IS417"
/db_xref="taxon:10598"

primer_bind 1..20

CDS

/gene="L1"
/note="My11 primer"
<1..>449
/gene="L1"
/note="My09/11"
/codon_start=1
/product="major capsid protein L1"
/protein_id="AAB19164.1"
/db_xref="GI:1209185"
/translation="AAGHNGICMGKQFLVYVDTTSTNWLCTEYTKNNNF
KRYVHREYDLOVFOCKITITFAVMTYHTNNSDILDMQFGALPPPSASLDY
REVTSQATTCOKTAPREKEDPLNKYTWENLEKESADLDQFPLGR"
1..449
/gene="L1"
complement(429..449)
primer_bind /note="My09 primer"

BASE COUNT 151 a 78 c 83 g 137 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 14; Length 449;
Best Local Similarity 95.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 20
|||||

RESULT 13
LOCUS 114275 22 bp DNA PAT 26-SEP-1995
DEFINITION Sequence 72 from patent US 5447839.
ACCESSION 114275
VERSION 114275.1 GI:997290
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Manos,M.Michele, Bauer,H.M., Greer,C.E., Resnick,R.M. and Ting,Y.
TITLE Detection of human papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5447839-A-72 05-SEP-1995;
FEATURES Location/Qualifiers
source 1..22
/organism="unknown"

BASE COUNT 9 a 3 c 7 g 3 t
ORIGIN

Query Match 90.0%; Score 18; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gcaactgaagtaactaag 20
|||||

Db 1 GCACCTGAGTAAGTAAG 18
RESULT 14
LOCUS 122732 22 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 220 from patent US 5527898.
ACCESSION 122732
VERSION 122732.1 GI:1603086
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Manos,M.Michele, Resnick,R.M.
and Zhang,T.Y.
TITLE Detection of human papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5527898-A-220 18-JUN-1996;
FEATURES Location/Qualifiers
source 1..22
/organism="unknown"

BASE COUNT 9 a 3 c 7 g 3 t
ORIGIN

Query Match 90.0%; Score 18; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gcaactgaagtaactaag 20
|||||

Db 1 GCACCTGAGTAAGTAAG 18

```

RESULT 15
LOCUS 147557 22 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 220 from patent US 5639871.
ACCESSION 147557
VERSION 147557.1 GI:2471522
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 22)
Bauer, H.M., Gravitt, P.E., Greer, C.E., Impraim, C.C.,
Manos, M., Michele, Resnick, R.M. and Zhang, T.Y.
TITLE Detection of human Papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5639871 A 220 17-JUN-1997;
FEATURES
Source 1..22
location/Qualifiers
BASE COUNT 9 a 3 c 7 g 3 t
ORIGIN

```

```

Query Match 90.0%; Score 18; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 gcactgaagtaactaagg 20
| | | | | | | | | | | | | | | | | | | | | |
DB 1 GCCTGAGTACTACTAGG 18

```

Search completed: April 3, 2002, 10:59:54
 Job time: 5185 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 09:33:29 ; Search time 321.05 Seconds
(without alignments)
53.408 Million cell updates/sec

Title: US-09-885-799-317

Perfect score: 20

Sequence: 1 atgacacgaactactaag 20

Scoring table:

IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues 1861242

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Maximum Match 100%

Listing first 45 summaries

Database :
N_Geneseq_1101: *
1: /SID2/gcgdata/geneseq/NA1980.DAT: *
2: /SID2/gcgdata/geneseq/NA1981.DAT: *
3: /SID2/gcgdata/geneseq/NA1982.DAT: *
4: /SID2/gcgdata/geneseq/NA1983.DAT: *
5: /SID2/gcgdata/geneseq/NA1984.DAT: *
6: /SID2/gcgdata/geneseq/NA1985.DAT: *
7: /SID2/gcgdata/geneseq/NA1986.DAT: *
8: /SID2/gcgdata/geneseq/NA1987.DAT: *
9: /SID2/gcgdata/geneseq/NA1988.DAT: *
10: /SID2/gcgdata/geneseq/NA1989.DAT: *
11: /SID2/gcgdata/geneseq/NA1990.DAT: *
12: /SID2/gcgdata/geneseq/NA1991.DAT: *
13: /SID2/gcgdata/geneseq/NA1992.DAT: *
14: /SID2/gcgdata/geneseq/NA1993.DAT: *
15: /SID2/gcgdata/geneseq/NA1994.DAT: *
16: /SID2/gcgdata/geneseq/NA1995.DAT: *
17: /SID2/gcgdata/geneseq/NA1996.DAT: *
18: /SID2/gcgdata/geneseq/NA1997.DAT: *
19: /SID2/gcgdata/geneseq/NA1998.DAT: *
20: /SID2/gcgdata/geneseq/NA1999.DAT: *
21: /SID2/gcgdata/geneseq/NA2000.DAT: *
22: /SID2/gcgdata/geneseq/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	20	100.0	30	Human papilloma vi
2	20	100.0	139	Target sequence in
3	19.6	98.0	449	Human papilloma vi
4	19	95.0	25	Human papilloma vi
5	19	95.0	25	Human papilloma vi
6	19	95.0	46	HPV linked oligonu
7	19	95.0	46	Human papilloma vi
8	19	95.0	22	Human papilloma vi
9	18	90.0	22	Human papilloma vi
10	18	90.0	22	Human papilloma vi
11	18	90.0	22	Human papilloma vi
			19	Probe MY93 for hum
			22	AAV17529

C 12	16.8	84.0	2233	22	AAF32525	Human male enhance
C 13	16.4	82.0	606	22	AAH87677	Peppermint plant o
C 14	16.4	82.0	2510	22	AAH32623	Human secreted pro
C 15	15.8	79.0	293	22	AAF68717	Human lung tumour
C 16	15.8	79.0	1953	21	AAQ99046	Human pancreatic c
C 17	15.8	79.0	2856	21	AAQ92291	Agarase 4-3 encodi
C 18	15.8	79.0	21278	21	AAA55965	Human G713 3'-end
C 19	15.2	76.0	286	21	AAQ3767	Human secreted exp
C 20	15.2	76.0	629	16	AAQ78996	Human immunoglobul
C 21	15.2	76.0	1121	21	AAQ36705	Arabidopsis thalija
C 22	15.2	76.0	1358	20	AAQ4324	Human secreted pro
C 23	15.2	76.0	1433	22	AAQ0142	Human cDNA clone 1
C 24	15.2	76.0	1506	21	AAQ77860	Human cancer assoc
C 25	15.2	76.0	2466	22	AAH32555	Human secreted pro
C 26	15.2	76.0	2636	22	AAH16363	Human cDNA sequenc
C 27	15.2	76.0	5956	22	AAH81804	Human differential
C 28	15.2	76.0	49999	20	AAZ23891	Murine LOBO genomi
C 29	15.2	76.0	49999	20	AAZ23891	Nucleotide sequenc
C 30	15	75.0	9192	12	AAQ11943	Fibrin clot bindin
C 31	14.8	74.0	40	17	AAQ70637	Fibrin binding clo
C 32	14.8	74.0	87	20	AAH85120	Human secreted pro
C 33	14.8	74.0	164	21	AAQ19108	Staphylococcus aur
C 34	14.8	74.0	203	18	AAV77446	Human secreted pro
C 35	14.8	74.0	327	21	AAQ05557	Human secreted pro
C 36	14.8	74.0	358	18	AAV78478	Staphylococcus aur
C 37	14.8	74.0	637	21	AAZ80402	Human colon cancer
C 38	14.8	74.0	1020	21	AAQ77870	Human cancer assoc
C 39	14.8	74.0	1712	22	AAH32537	Human secreted pro
C 40	14.8	74.0	9468	19	AAV25602	Feline immunodefici
C 41	14.8	74.0	13321	18	AAV4315	Human HCR DNA. Ho
C 42	14.8	74.0	14483	22	AAH45311	Polymorphic repeat
C 43	14.8	74.0	50000	21	AAH90366	Human yeast gene.
C 44	14.8	74.0	119950	20	AAH90201	Mycoplasma genitali
C 45	14.8	74.0	580073	18	AAH58840	

ALIGNMENTS

RESULT 1	AAAT00997	standard; DNA; 30 BP.
AC	AAAT00997:	
AT	02-Apr-1996 (first entry)	
DE	Human papilloma virus HPV-58-specific probe.	
DE	Human papilloma virus; HPV-58; cervical carcinoma; screening; PCR;	
KW	elongated general primer; nucleic acid amplification; probe: ss.	
KW	Synthetic.	
OS	W09522626-A1.	
XX	24-AUG-1995.	
XX	20-FEB-1995.	95MO-NL00066.
XX	23-SEP-1994.	94EP-0202739.
XX	21-FEB-1994.	94EP-0200432.
XX	(RESE-) STICHTING RESEARCHFONDS PATHOLOGIE.	
XX	Meijer GJM, Snijders PJF, Van Den Brule AJC, Walboomers JMM;	
XX	WPI; 1995-302728/39.	
XX	Elongated general primer(s) and Human Papilloma Virus-specific	
XX	probe(s) - for use in amplification and detection method(s)	
XX	providing improved HPV detection in cervical smear(s).	

PS Claim 29; Page 57; 61pp; English.
 XX
 CC Elongated general primers GP5+ and GP6+ and their subseq. derivs.
 CC are used for amplifying Human Papilloma Virus sequences. Unlike the
 CC known general primers GP5 and GP6, originally selected from the HPV
 CC L1 region on the basis of sequence information of HPV-6, -16, -18,
 CC -31 and -33, the 3'-elongated primers overcome reduced PCR efficiency
 CC and can distinguish HPV status in cytomorphologically normal cervical
 CC smears which previously gave ambiguous or negative results.
 CC The present sequence is that of a probe used for detecting HPV-58
 CC sequences among PCR products amplified by the elongated primers.
 XX
 SQ Sequence 30 BP; 12 A; 4 C; 7 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 16; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.31;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgacatgaagtaactaagg 20
 ||||||||||||||||||
 DB 4 atgacatgaagtaactaagg 23

RESULT 2
 AAV37221
 ID AAV37221 standard; DNA; 139 BP.
 XX
 AC AAV37221;
 XX
 DT 14-SEP-1998 (first entry)
 XX
 DE Target sequence in the L1 gene of HPV type 58.
 XX
 KM h1 gene; HPV type 58; hybridisation assay; detection; HPV type;
 KM Cervical cancer; ss.
 KM
 OS Human papillomavirus.
 XX
 PN W09817829-A2.
 XX
 PD 30-APR-1998.
 XX
 PE 17-OCT-1997; 97WO-US19467.
 XX
 PR 25-OCT-1996; 96US-0739103.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Abravaya K, Gorzowski JJ, Hoehnle RJ, Kroeger PE;
 PI Moore JJ;
 XX
 DR WPI; 1998-261520/23.
 XX
 PT Detection of oncogenic human papilloma virus types - using
 PT oligonucleotide probes which hybridise with a 140 bp region of the
 PT L1 gene found in the human papillomavirus genome
 XX
 PS Disclosure; Page 27; 33pp; English.
 XX
 CC The present sequence represents a portion of the L1 gene of oncogenic
 CC human papillomavirus (HPV) type 58. Oligonucleotide AAV38493 is used as
 CC a primer and probe to detect this organism. The probes of the invention
 CC can be used in a hybridisation assay to detect HPV types. They can also
 CC be used in a similar way with the addition that HPV target sequences
 CC can be amplified first. The products can be used for the detection of
 CC HPV types, particularly those linked to the development of cervical
 CC cancer.
 XX
 SQ Sequence 139 BP; 47 A; 18 C; 26 G; 48 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 139;

Best Local Similarity 100.0%; Pred. No. 0.37;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgacatgaagtaactaagg 20
 ||||||||||||||||||
 DB 45 atgacatgaagtaactaagg 64

RESULT 3
 AAT91255
 ID AAT91255 standard; DNA; 449 BP.
 XX
 AC AAT91255;
 XX
 DT 24-APR-1998 (first entry)
 XX
 DE Human papilloma virus HPV-58 DNA sequence for generating probes.
 XX
 KM Human papilloma virus; HPV; probe; isolation; detection; hybridise; ds.
 XX
 OS Human papillomavirus.
 XX
 FH Key Location/Qualifiers
 FT 1..53
 FT /tag= a
 FT /note= "claimed probe for HPV-58"
 FT misc-feature
 FT 55..104
 FT /tag= b
 FT /note= "claimed probe for HPV-58"
 FT 125..210
 FT /tag= c
 FT /note= "claimed probe for HPV-58"
 FT 212..234
 FT /tag= d
 FT /note= "claimed probe for HPV-58"
 FT 242..293
 FT /tag= e
 FT /note= "claimed probe for HPV-58"
 FT 295..428
 FT /tag= f
 FT /note= "claimed probe for HPV-58"
 FT 430..449
 FT /tag= g
 FT /note= "claimed probe for HPV-58"

W09739010-A1.

23-OCT-1997.

14-APR-1997; 97WO-US06354.

15-APR-1996; 96US-0015427.

(UYNE-) UNIV NEW MEXICO STATE.

Manos MM, Wheeler CM;

WPI; 1997-526386/48.

Nucleic acid probes for human papilloma virus - for isolation and
 detection of homologous or variant regions of DNA from different
 viruses

Claim 8; Page 26; 48pp; English.

The present sequence represents a DNA sequence specific to a human
 papilloma virus (HPV) which is used to generate specifically claimed
 probes (see features table). The probes can be used to hybridise to
 the homologous regions of the specific HPV. The probes can also be
 used to generate probes with nucleotide substitutions, which will
 hybridise to newly discovered variant regions of HPV. The probes are
 useful for the isolation and detection of HPV nucleic acid sequences.

Sequence 449 BP; 145 A; 78 C; 79 G; 137 T; 10 other;

Query Match 99.0%; Score 19.6; DB 18; Length 449;
Best Local Similarity 95.0%; Pred. No. 0.69;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 20
|||||
DB 87 atgcactgaagtaactaag 106

RESULT 4

ID AAT44683 standard; DNA; 25 BP.

AC AAT44683;

DT 28-JAN-1997 (first entry)

DE Human papillomavirus detection probe MY179 for HPV type 58.

KW Probe: primer; PCR; polymerase chain reaction; amplification;
human papillomavirus; consensus; ss.

OS Synthetic.

PN US5527898-A.

PD 18-JUN-1996.

PF 09-SEP-1988; 88US-0243486.

PR 24-SEP-1993; 93US-0126452.

PR 09-SEP-1988; 88US-0243486.

PR 10-MAR-1989; 89US-0322550.

PR 09-SEP-1989; 89MO-US03747.

PR 14-NOV-1990; 90US-0613142.

PR 20-APR-1993; 93US-0050743.

PR 07-JUN-1995; 95US-0474542.

PA (HOFF) HOFFMANN LA ROCHE INC.

PI Bauer HM, Gravitt PE, Greer CE, Manos MM, Resnick RM;

PI Zhang TY;

DR WPI; 1996-245903/30.

XX Nucleic acid hybridisation probes - specific for selected human

PT papilloma virus types

PS Claim 1; Column 85; 96pp; English.

CC The invention relates to new oligonucleotide probes and primers used

CC for the detection of human papillomaviruses which are not genital types

CC 6, 11, 16, 18 or 33. The probes and primers AAT4468-T44693 are esp.

CC Used to detect HPV types 26, 31, 31B, 35, 39, 40, 43, 45, 51-59 and 68.

CC The primers can be used to detect these HPV types in conjunction with

CC the consensus primers and typing probes AAT4473-T44906, which are based

CC on and amplify fragments of the L1, E6, E7 and E1 regions of the HPV

CC sequences. This primer is targeted to the new HPV type 58.

XX Sequence 25 BP; 10 A; 4 C; 5 G; 6 T; 0 other;

Query Match 95.0%; Score 19; DB 17; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.99;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 19
|||||
DB 7 atgcactgaagtaactaag 25

RESULT 5

ID AAT77923 standard; DNA; 25 BP.

AC AAT77923;

DT 06-OCT-1997 (first entry)

DE Papillomavirus 251 probe MYB179.

KW Papillomavirus 251; PAP251; probe; detection; ss.

OS Synthetic.

PN US5639871-A.

PD 17-JUN-1997.

PF 09-SEP-1988; 88US-0243486.

PR 24-SEP-1993; 93US-0126452.

PR 09-SEP-1988; 88US-0243486.

PR 10-MAR-1989; 89US-0322550.

PR 09-SEP-1989; 89MO-US03747.

PR 14-NOV-1990; 90US-0613142.

PR 20-APR-1993; 93US-0050743.

PR 01-JUN-1995; 95US-0457648.

PA (HOFF) ROCHE MOLECULAR SYSTEMS INC.

PI Bauer HM, Gravitt PE, Greer CE, Imprim CC, Manos MM;

PI Resnick RM, Zhang TY;

DR WPI; 1997-332084/30.

XX New oligonucleotide probes for human papilloma-virus - used for

PT detecting and typing HPV and for detecting previously unknown HPV

PS types and subtypes

PS Disclosure; Columns 83-84; 94pp; English.

XX The present sequence is a papillomavirus 251 (PAP251) specific

CC probe.

XX Sequence 25 BP; 10 A; 4 C; 5 G; 6 T; 0 other;

Query Match 95.0%; Score 19; DB 18; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.99;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 19
|||||
DB 7 atgcactgaagtaactaag 25

RESULT 6

ID AAT44888 standard; DNA; 46 BP.

AC AAT44888;

DT 05-FEB-1997 (first entry)

DE HPV linked oligonucleotide probe TY2018.

KW Probe: primer; PCR; polymerase chain reaction; amplification;

OS human papillomavirus; consensus; ss.

PN US5527898-A.

XX 18-JUN-1996.
 PD 09-SEP-1988; 88US-0243486.
 XX 24-SEP-1993; 93US-0126452.
 PR 09-SEP-1988; 88US-0243486.
 PR 10-MAR-1989; 89US-0322550.
 PR 09-SEP-1989; 89MO-US03747.
 PR 14-NOV-1990; 90US-0613142.
 PR 20-APR-1993; 93US-0050743.
 PR 07-JUN-1995; 95US-0475442.

(HOFF) HOFFMANN LA ROCHE INC.

XX Bauer HM, Gravitt PE, Greer CE, Manos MM, Resnick RM,
 PI Zhang TY;
 XX WPI; 1996-299903/30.

XX Nucleic acid hybridisation probes - specific for selected human
 PT papilloma virus types

XX Example 7; Column 59-60; 96pp; English.

CC The invention relates to new oligonucleotide probes and primers used
 CC for the detection of human papillomaviruses (HPV) which are not genital
 CC types 6, 11, 16, 18 or 33. The probes and primers AAT44608-T44693 are
 CC esp. used to detect HPV types 26, 31, 31B, 35, 39, 40, 43, 45, 51-59 and
 CC 68. The primers can be used to detect these HPV types in conjunction with
 CC the consensus primers and typing probes AAT44733-T44906, which are based
 CC on and amplify fragments of the L1, E6, E7 and E1 regions of the HPV
 CC sequences. Detection of the amplification prods. is done with probes
 CC derived from consensus sequences found in all characterised HPV
 CC sequences.
 CC Probes AAT44874-89 are probes used to detect amplified prods. from the
 CC L1 region. Each probe contains two hybridisation sequences specific to
 CC each of 2 hypervariable region of the L1 sequence. This results in a
 CC doubling of probes available for detection by an immobilisation assay in
 CC a microtitre plate. This probe is specific for HPV type 58.

XX Sequence 46 BP; 20 A; 10 C; 9 G; 7 T; 0 other;

Query Match 95.0%; Score 19; DB 17; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaa 19
 Db 28 atgcactgaagtaactaa 46

RESULT 7

AAT78109 standard; DNA; 46 BP.

XX AAT78109;

XX 07-OCT-1997 (first entry)

XX Human papillomavirus 58 linked oligonucleotide probe TY2018.

KM Human: Papillomavirus 58; HPV58; linked oligonucleotide probe;

KW detection; ss.

OS Synthetic.

XX US5639871-A.

XX 17-JUN-1997.

PF 09-SEP-1988; 88US-0243486.

XX 24-SEP-1993; 93US-0126452.
 PR 09-SEP-1988; 88US-0243486.
 PR 10-MAR-1989; 89US-0322550.
 PR 09-SEP-1989; 89MO-US03747.
 PR 14-NOV-1990; 90US-0613142.
 PR 20-APR-1993; 93US-0050743.
 PR 01-JUN-1995; 95US-0457648.

(HOFF) ROCHE MOLECULAR SYSTEMS INC.

XX Bauer HM, Gravitt PE, Greer CE, Imprim CC, Manos MM;
 PI Resnick RM, Zhang TY;

XX WPI; 1997-332084/30.

XX New oligo-nucleotide probes for human papilloma-virus - used for
 PT detecting and typing HPV and for detecting previously unknown HPV
 PT types and subtypes

XX Example 7; Columns 153-154; 94pp; English.

CC The present sequence is a human papillomavirus 58 (HPV58)
 CC linked oligonucleotide probe.

XX Sequence 46 BP; 20 A; 10 C; 9 G; 7 T; 0 other;

Query Match 95.0%; Score 19; DB 18; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaa 19
 Db 28 atgcactgaagtaactaa 46

RESULT 8

AAQ98646 standard; DNA; 22 BP.

XX AAQ98646;

XX 10-APR-1996 (first entry)

XX Human papilloma virus PAP251 specific oligonucleotide probe MY93.

KW Human papilloma virus; probe; detection; diagnosis; genital;

KW oral; carcinomas; research; typing; PAP251; specific; MY93; ss.

OS Synthetic.

XX US5447839-A.

XX 05-SEP-1995.

XX 09-SEP-1988; 88US-0243486.

XX 14-NOV-1990; 90US-0613142.

XX 09-SEP-1988; 88US-0243486.

XX 10-MAR-1989; 89US-0322550.

XX 09-SEP-1989; 89MO-US03747.

XX 20-APR-1993; 93US-0050743.

(HOFF) HOFFMANN LA ROCHE INC.

XX Bauer HM, Greer CE, Manos MM, Resnick RM, Ting Y;

XX WPI; 1995-319884/41.

XX Detection of human papilloma virus DNA by amplification - using
 PT specific consensus primer pairs and pred. detection with generic or
 PT type specific probes for use in research and diagnosis

XX Claim 3; Columns 61-62; 36pp; English.
 PS
 CC The human papilloma virus (HPV) specific probes AAG98564-098650 are
 CC used to detect, or type HPV for research or diagnostic purposes,
 CC e.g. to identify HPV that are implicated in genital or oral
 CC carcinomas.
 CC
 SQ Sequence 22 BP; 9 A; 3 C; 7 G; 3 T; 0 other;

Query Match 90.0%; Score 18; DB 16; Length 22;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 gcaactgaagtaactaagg 20
 Db 1 gcaactgaagtaactaagg 18

RESULT 9

AAT44681
 ID AAT44681 standard; DNA: 22 BP.

AC AAT44681;

DT 28-JAN-1997 (first entry)

DE Human papillomavirus detection probe MY93 for HPV type 58.

KW Probe; primer; PCR; polymerase chain reaction; amplification;

KM human papillomavirus; consensus; ss.

OS Synthetic.

PN US5527898-A.

PD 18-JUN-1996.

PF 09-SEP-1988; 88US-0243486.

PR 24-SEP-1993; 93US-0126452.

PR 09-SEP-1988; 88US-0243486.

PR 10-MAR-1989; 89US-0322550.

PR 09-SEP-1989; 89WO-US03747.

PR 14-NOV-1990; 90US-0613142.

PR 20-APR-1993; 93US-0050743.

PR 07-JUN-1995; 95US-0474542.

PA (HOFF) HOFFMANN LA ROCHE INC.

PI Bauer HM, Gravitt PE, Greer CE, Manos MM, Resnick RM;

PI Zhang TY;

DR WPI: 1996-299903/30.

XX Nucleic acid hybridisation probes - specific for selected human

XX papilloma virus types

XX Claim 1; Column 147; 96pp; English.

XX The invention relates to new oligonucleotide probes and primers used

XX for the detection of human papillomaviruses which are not genital types

XX 6, 11, 16, 18 or 33. The probes and primers AAT44608-744693 are esp.

XX used to detect HPV types 26, 31, 31B, 35, 39, 40, 43, 45, 51-59 and 68.

XX The primers can be used to detect these HPV types in conjunction with

XX the consensus primers and typing probes AAT44733-744906, which are based

XX on and amplify fragments of the L1, B6, E7 and E1 regions of the HPV

XX sequences. This primer is targeted to the new HPV type 58.

XX Sequence 22 BP; 9 A; 3 C; 7 G; 3 T; 0 other;

Query Match 90.0%; Score 18; DB 17; Length 22;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 gcaactgaagtaactaagg 20
 Db 1 gcaactgaagtaactaagg 18

RESULT 10

AA78084
 ID AA78084 standard; DNA: 22 BP.

AC AA78084;

DT 07-OCT-1997 (first entry)

DE Human papillomavirus 58 specific typing probe MY93.

KW Human; papillomavirus 58; HPV58; typing probe;

KM detection; ss.

OS Synthetic.

PN US5639871-A.

PD 17-JUN-1997.

PF 09-SEP-1988; 88US-0243486.

PR 24-SEP-1993; 93US-0126452.

PR 09-SEP-1988; 88US-0243486.

PR 10-MAR-1989; 89US-0322550.

PR 09-SEP-1989; 89WO-US03747.

PR 14-NOV-1990; 90US-0613142.

PR 20-APR-1993; 93US-0050743.

PR 01-JUN-1995; 95US-0457648.

PA (HOFF) ROCHE MOLECULAR SYSTEMS INC.

PI Bauer HM, Gravitt PE, Greer CE, Imprim CC, Manos MM;

PI Resnick RM, Zhang TY;

DR WPI: 1997-332084/30.

XX New oligo:nucleotide probes for human papilloma-virus - used for

XX detecting and typing HPV and for detecting previously unknown HPV

XX types and subtypes

XX Disclosure; Columns 145-146; 94pp; English.

XX The present sequence is a human papillomavirus 58 (HPV58)

XX specific typing probe.

XX Sequence 22 BP; 9 A; 3 C; 7 G; 3 T; 0 other;

XX Query Match 90.0%; Score 18; DB 18; Length 22;

XX Best Local Similarity 100.0%; Pred. No. 3.2;

XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OY 3 gcaactgaagtaactaagg 20

XX Db 1 gcaactgaagtaactaagg 18

XX RESULT 11

XX AAV17529

XX ID AAV17529 standard; DNA: 22 BP.

XX AC AAV17529;

XX DT 04-JUN-1998 (first entry)

```

XX DE Probe MY93 for human papillomavirus typing.
XX KW Human papillomavirus; HPV; HPV detection; HPV typing;
XX KW LI type-specific probe; ss.
XX OS Synthetic.
XX OS Human papillomavirus.
XX PN US705627-A.
XX PD 06-JAN-1998.
XX PF 26-MAY-1995; 95US-0452055.
XX PR 09-SEP-1989; 89US-0613142.
XX PR 09-SEP-1988; 88US-0243486.
XX PR 10-MAR-1989; 89US-0322550.
XX PR 20-APR-1993; 93US-0050743.
XX PR 26-MAY-1995; 95US-0452055.
XX PA (HOF) ROCHE MOLECULAR SYSTEMS INC.
XX PI Bauer HM, Greer CE, Manos MM, Resnick RM, Ting Y;
XX DR WPI; 1998-192210/17.
XX PT Human papilloma probes and primers - useful for, e.g. detecting and
XX PT typing of human papilloma viruses
XX PS Disclosure; Column 17-18; 37pp; English.
XX CC This sequence represents a human papillomavirus (HPV) L1 type-specific
XX CC probe of the invention. This sequence may be used in conjunction with L1
XX CC specific primers for detecting and typing HPV. Identification and typing
XX CC of HPV is important as different types of HPV pose different risks for
XX CC infected individuals. HPV16 and HPV18 have been more consistently
XX CC identified in higher grades of cervical dysplasia and carcinoma than
XX CC other HPV types.
XX SQ Sequence 22 BP; 9 A; 3 C; 7 G; 3 T; 0 other;

Query Match 90.0%; Score 18; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 gcactgaactaactaag 20
DB 1 gcactgaactaactaag 18

RESULT 12
AAE32525/C
ID AAE32525 standard; cDNA; 2233 BP.
XX AC AAE32525;
XX DE 19-APR-2001 (first entry)
XX DE Human male enhanced antigen-2 (MEA-2) nucleotide sequence SEQ ID NO:19.
XX KW Human; male enhanced antigen-2; MEA-2; identification; spermatogenesis;
XX KW spermatogenesis disease; chromosome marker; pancreatic cancer; ss.
XX OS Homo sapiens.
XX PN JP2000316580-A.
XX PD 21-NOV-2000.
XX PF 30-APR-1999; 99JP-0125196.
XX

```

```

PR 30-APR-1999; 99JP-0125196.
XX PA (ITO) ITO HAM KK.
XX DR WPI; 2001-128256/14.
XX DR P-PSDB: AAB69071.
XX PT A new protein, human male-enhanced antigen-2, useful for detecting
XX PT spermatogenesis diseases
XX PS Claim 6; Page 17-18; 21pp; Japanese.
XX CC The present sequence encodes the human male enhanced antigen-2 (MEA-2).
XX CC The present invention also described an antibody specific for the
XX CC MEA-2 protein. The antibody can be used for the identification of a
XX CC gene causing diseases related to spermatogenesis. The MEA-2 nucleotide
XX CC sequence is useful as a chromosome marker, and in the detection of
XX CC pancreatic cancer.
XX SQ Sequence 2233 BP; 506 A; 543 C; 580 G; 604 T; 0 other;

```

```

Query Match 84.0%; Score 16.8; DB 22; Length 2233;
Best Local Similarity 90.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 atgacatgaactaactaag 20
DB 1292 ATGCACTGAAGTAAGTAAGTCTGG 1273

```

```

RESULT 13
AAH87677
ID AAH87677 standard; cDNA; 606 BP.
XX AC AAH87677;
XX DE 25-SEP-2001 (first entry)
XX DE Peppermint plant oil gland expressed cDNA 33.
XX KW Peppermint; plant oil gland cell; terpenoid essential oil; resin;
XX KW genetic mapping; antisense suppression; recombinant expression; ss.
XX OS Mentha x piperita.
XX PN WO200153319-A1.
XX PD 26-JUL-2001.
XX PF 19-JAN-2001; 2001WO-0502567.
XX PR 20-JAN-2000; 2000US-0177264.
XX PA (CROT/) CROTEAU R B.
XX PA (LANG/) LANGE B M.
XX PA (WILD/) WILDUNG M R.
XX PI Croteau RB, Lange BM, Wildung MR;
XX DR WPI; 2001-488706/53.
XX PT New nucleic acid molecules corresponding to mRNA molecules expressed in
XX PT peppermint oil glands for enhancing expression of plant oil gland cell
XX PT proteins -
XX PS Claim 1; Page 89; 251pp; English.
XX CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
XX CC correspond to all or part of a mRNA molecule expressed in plant oil
XX CC gland cells, especially peppermint and plant oil glands that produce
XX CC terpenoid essential oils and resins. The nucleic acids are useful for
XX CC genetically mapping a plant genome for genes expressed in plant oil

```

CC gland cells and to suppress (for example by antisense suppression) or
 CC enhance their expression (for example by genetically transforming a
 CC plant cell with a replicable expression vector that expresses one or more
 CC proteins naturally expressed in plant oil gland cells). The nucleic acids
 CC are also useful for recombinant expression of plant oil gland proteins
 CC required for terpenoid essential oil and/or resin production in bacterial
 CC and/or yeast cells.
 CC
 SO Sequence 606 BP; 168 A; 117 C; 148 G; 172 T; 1 other;
 Query Match 82.0%; Score 16.4; DB 22; Length 606;
 Best Local Similarity 94.4%; Pred. No. 31;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 gcaatgaaglaactaag 20
 Db 110 gcaatgaaglaactaag 127
 ||| ||||| ||||| |||||
 RESULT 14
 ID AAH32623/c
 AC AAH32623 standard; cDNA; 2510 BP.
 XX
 AC AAH32623;
 XX
 DT 10-AUG-2001 (first entry)
 DE
 XX Human secreted protein gene 34 cDNA clone HBXCT44, SEQ ID NO:112.
 XX
 KW Human; secreted protein; proliferative disorder; cancer; chromosome 9;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder; tumour;
 KW endocrine disorder; infection; wound healing; vulnerability;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO200134628-A1.
 XX
 PD 17-MAY-2001.
 PD
 XX 08-NOV-2000; 2000WO-US30653.
 PF
 XX 12-NOV-1999; 99US-0164735.
 PR
 XX 27-JUL-2000; 2000US-0221193.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 PI Ruben SM, Komatsoulis GA, Blurse CE, Ni J, Moore PA;
 XX
 DR WPI: 2001-329066/34.
 DR P-PSDB: AAG73444.
 DR
 XX
 PT Nucleic acids encoding 35 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 PT
 XX
 PS Claim 4; Page 512-513; 604pp; English.
 XX
 CC AAH32622-AAH32627 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73446-AAG73448 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of

CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunoassay (ELISA). The present sequence represents a human
 CC secreted protein-encoding cDNA of the invention.
 CC
 SO Sequence 2510 BP; 563 A; 683 C; 635 G; 623 T; 6 other;
 Query Match 82.0%; Score 16.4; DB 22; Length 2510;
 Best Local Similarity 85.0%; Pred. No. 37;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 atgcactgaaglaactaag 20
 Db 1459 ATTCACTGCAGTAAGTARG 1440
 ||| ||||| ||||| |||||
 RESULT 15
 ID AAF68717 standard; cDNA; 293 BP.
 XX
 AC AAF68717;
 XX
 DT 12-APR-2001 (first entry)
 DE
 XX Human lung tumour protein related nucleotide sequence SEQ ID NO:652.
 DE
 XX Human lung tumour; lung tumour protein; gene therapy;
 KW Lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
 KW cytostatic; antisense inhibition; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO200100828-A2.
 XX
 PD 04-JAN-2001.
 PD
 XX 30-JUN-2000; 2000WO-US18061.
 PF
 XX 30-JUN-1999; 99US-0346492.
 PR
 XX 15-OCT-1999; 99US-0419356.
 PR
 XX 17-DEC-1999; 99US-0466867.
 PR
 XX 30-DEC-1999; 99US-0476300.
 PR
 XX 06-MAR-2000; 2000US-0519642.
 PR
 XX 22-MAR-2000; 2000US-0533077.
 PR
 XX 10-APR-2000; 2000US-0546259.
 PR
 XX 27-APR-2000; 2000US-0560406.
 PR
 XX 05-JUN-2000; 2000US-0589184.
 XX
 PA (CORI-) CORIAX CORP.
 PA
 XX Wang T, Banour CS, Lodes MJ, Fanger GR, Vedrick TS, Carter D;
 PI Reltter MW, Mannion J;
 XX
 DR WPI: 2001-071488/08.

XX Lung tumor-associated proteins and the nucleic acids that encode them,
 PT useful for preventing, diagnosing and treating lung cancer -
 XX
 PS Claim 4; Page 365; 436pp; English.

CC The present invention describes immunogenic portions of lung tumour-
 CC associated proteins (I) and the nucleic acids (NAs) that encode them.
 CC (i) have cytostatic activity and can be used in gene therapy, antisense
 CC inhibition and in vaccines. The NAs and the lung tumour-associated
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with their inappropriate expression,
 CC especially lung cancers. For example, the NAs may be administered to
 CC treat diseases by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of the protein by expressing inactive proteins
 CC or to supplement the patient's own production of (I). Additionally, the
 CC NAs may be used to produce the lung-tumour associated protein, according
 CC to standard recombinant DNA methodology. Conversely, antisense NA
 CC molecules may be administered to down regulate protein expression by
 CC binding with the cells own genes and preventing their expression. The NA
 CC and complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar NA sequences in
 CC samples, and hence which patients may be in need of treatment for lung
 CC cancer. The (I) may be used as antigens in the production of antibodies
 CC and in assays to identify modulators (agonists and antagonists) of the
 CC expression and activity of the protein. AAF68083 to AAF68878 and
 CC AAB76848 to AAB76878 represent human lung tumour protein related
 CC nucleotide and protein sequences which are used in the exemplification
 CC of the present invention.
 XX

SQ Sequence 293 BP; 103 A; 60 C; 61 G; 67 T; 2 other;

Query Match 79.08; Score 15.8; DB 22; Length 293;
 Best Local Similarity 89.58; Pred. No. 58;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 tgactgaagtactaag 20
 ||||| |||||
 Db 8 tgactgaagtactaag 26

Search completed: April 3, 2002, 11:05:27
 Job time: 5518 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 09:33:29 ; Search time 143.87 Seconds
(without alignments)
31.484 Million cell updates/sec

Title: US-09-885-799-317

Perfect Score: 20

Sequence: 1 atgcactgaagtaactaaga 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	139	4	US-08-739-103A-33
2	19	95.0	25	1	US-08-474-542A-59
3	19	95.0	25	1	US-08-457-648-59
4	19	95.0	46	1	US-08-474-542A-245
5	19	95.0	46	1	US-08-457-648-245
6	18	90.0	22	1	US-08-050-743-72
7	18	90.0	22	1	US-08-474-542A-220
8	18	90.0	22	1	US-08-457-648-220
9	18	90.0	22	1	US-08-452-055-72
10	15.2	76.0	629	3	US-08-545-809A-58
11	14.8	74.0	87	1	US-08-433-126A-63
12	14.8	74.0	87	1	US-08-433-126A-63
13	14.8	74.0	87	3	US-08-976-413A-63
14	14.8	74.0	87	5	PCT-US96-06059-63
15	14.8	74.0	637	4	US-09-328-111-486
16	14.8	74.0	9468	1	US-08-325-947-10
17	14.4	72.0	3666	2	US-08-682-517-13
18	14.4	72.0	3666	2	US-08-682-517-13
19	14.4	72.0	4197	2	US-08-682-517-14
20	14.4	72.0	4197	2	US-08-682-517-7
21	14.4	72.0	4643	2	US-08-682-517-8
22	14.4	72.0	5578	1	US-08-605-106-6
23	14.4	72.0	176373	3	US-08-081-610-2
24	14.2	71.0	1713	4	US-09-128-155-17
25	14.2	71.0	1734	3	US-09-347-798-7
26	14.2	71.0	1734	3	US-08-858-052-2
27	14.2	71.0	1892	4	US-09-200-284-2
					US-09-423-340-3

28	14.2	71.0	11288	4	US-08-646-301A-1	Sequence 1, Appl
29	14.2	71.0	11288	4	US-08-481-968A-4	Sequence 4, Appl
30	14.2	71.0	87350	3	US-08-781-891-79	Sequence 79, Appl
31	13.8	69.0	19	4	US-08-739-103A-13	Sequence 13, Appl
32	13.8	69.0	19	4	US-08-739-103A-14	Sequence 14, Appl
33	13.8	69.0	139	4	US-08-739-103A-12	Sequence 12, Appl
34	13.8	69.0	1527	4	US-07-956-483-26	Sequence 26, Appl
35	13.8	69.0	1527	4	US-08-472-240A-19	Sequence 19, Appl
36	13.8	69.0	6914	1	US-08-920-812-22	Sequence 22, Appl
37	13.8	69.0	6914	1	US-08-920-827-22	Sequence 22, Appl
38	13.8	69.0	6914	1	US-08-921-177-22	Sequence 22, Appl
39	13.8	69.0	6914	1	US-08-362-577C-22	Sequence 22, Appl
40	13.8	69.0	6914	2	US-08-920-828-22	Sequence 22, Appl
41	13.8	69.0	9737	2	US-08-944-449-7	Sequence 7, Appl
42	13.6	68.0	195	4	US-08-905-223-106	Sequence 106, Appl
43	13.6	68.0	1957	4	US-09-352-990-11	Sequence 11, Appl
44	13.6	68.0	2085	2	US-08-668-128B-7	Sequence 7, Appl
45	13.6	68.0	2085	2	US-08-905-445-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-739-103A-33
Sequence 33, Application US/08739103A
Patent No. 6265154
GENERAL INFORMATION:
APPLICANT: P. Kroeger
APPLICANT: K. Abiravaya
APPLICANT: J. Gorzowski
APPLICANT: R. Moore
APPLICANT: J. Moore
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FOR DETECTING
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.0.1
SOFTWARE: Microsoft word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/739,103A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Paul D. Yaeger
REGISTRATION NUMBER: 37,477
REFERENCE/DOCKET NUMBER: 5990.US.O1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-2341
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA (HPV type 58)
US-08-739-103A-33

Query Match 100.0%; Score 20; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 20
|||||
DB 45 ATGCACCTGAAGTAACTAAG 64

RESULT 2

US-08-474-542A-59
; Sequence 59, Application US/08474542A
; Patent No. 5527896

GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravitt, Patti E.
APPLICANT: Greer, Catherine E.
APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the
TITLE OF INVENTION: Polymerase Chain Reaction
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,542A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9234
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-542A-59

Query Match 95.0%; Score 19; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgcactgaagtaactaag 19
|||||
DB 7 ATGCACCTGAAGTAACTAAG 25

RESULT 3

US-08-457-648-59
; Sequence 59, Application US/08457648
; Patent No. 5639871

GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravitt, Patti E.
APPLICANT: Greer, Catherine E.
APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the

TITLE OF INVENTION: Polymerase Chain Reaction
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,648
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-457-648-59

Query Match 95.0%; Score 19; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgcactgaagtaactaag 19
|||||
DB 7 ATGCACCTGAAGTAACTAAG 25

RESULT 4

US-08-474-542A-245
; Sequence 245, Application US/08474542A
; Patent No. 5527896

GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravitt, Patti E.
APPLICANT: Greer, Catherine E.
APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the
TITLE OF INVENTION: Polymerase Chain Reaction
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,542A
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9234
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ. ID NO: 245:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-542A-245

Query Match 95.0%; Score 19; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 19
|||||
DB 28 ATGCACGTGAAGTAACTAAG 46

RESULT 5
US-08-457-648-245
Sequence 245, Application US/08457648
Patent No. 5639871
GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravitt, Patti E.
APPLICANT: Greer, Catherine E.
APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,648
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ. ID NO: 245:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-457-648-245

Query Match 95.0%; Score 19; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaag 19
|||||
DB 28 ATGCACGTGAAGTAACTAAG 46

RESULT 6
US-08-050-743-72
Sequence 72, Application US/08050743
Patent No. 5447839
GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Greer, Catherine E.
APPLICANT: Manos, Michele
APPLICANT: Resnick, Robert M.
APPLICANT: Yimg, Yi
TITLE OF INVENTION: Detection of Human Papillomavirus by the
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/050,743
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sias, Stacey R.
REGISTRATION NUMBER: 32,630
REFERENCE/DOCKET NUMBER: 8793
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2863
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ. ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-050-743-72

Query Match 90.0%; Score 18; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gcactgaagtaactaag 20
|||||
DB 1 GCACGTGAAGTAACTAAG 18

RESULT 7
US-08-474-542A-220
Sequence 220, Application US/08474542A
Patent No. 5527898
GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravitt, Patti E.
APPLICANT: Greer, Catherine E.


```

? APPLICANT: Impraim, Chaka C.
? APPLICANT: Manos, M. Michele
? APPLICANT: Resnick, Robert M.
? TITLE OF INVENTION: Detection of Human Papillomavirus by the
? TITLE OF INVENTION: Polymerase Chain Reaction
? NUMBER OF SEQUENCES: 298
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hoffmann-La Roche Inc.
? STREET: 340 Kingsland Street
? CITY: Nutley
? STATE: New Jersey
? COUNTRY: U.S.A.
? ZIP: 07110
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/474,542A
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Petry, Douglas A.
? REGISTRATION NUMBER: 35,321
? REFERENCE/DOCKET NUMBER: 9234
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (510) 814-2974
? TELEFAX: (510) 814-2977
? INFORMATION FOR SEQ ID NO: 220:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 22 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-474-542A-220

Query Match          90.0%; Score 18; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gcactgaagtaactaagg 20
DB 1 GCACGTGAAGTAACTAAGG 18

RESULT 8
? Sequence 220, Application US/08457648
? Patent No. 5639871
? GENERAL INFORMATION:
? APPLICANT: Bauer, Heidi M.
? APPLICANT: Greer, Catherine E.
? APPLICANT: Manos, M. Michele
? APPLICANT: Impraim, Chaka C.
? APPLICANT: Resnick, Robert M.
? TITLE OF INVENTION: Detection of Human Papillomavirus by the
? TITLE OF INVENTION: Polymerase Chain Reaction
? NUMBER OF SEQUENCES: 298
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hoffmann-La Roche Inc.
? STREET: 340 Kingsland Street
? CITY: Nutley
? STATE: New Jersey
? COUNTRY: U.S.A.
? ZIP: 07110
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/452,055
? FILING DATE:
? CLASSIFICATION: 536
? ATTORNEY/AGENT INFORMATION:
? NAME: Stias, Stacey R.
? REGISTRATION NUMBER: 32,630
? REFERENCE/DOCKET NUMBER: 9188
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (510) 814-2863
? TELEFAX: (510) 814-2977
? INFORMATION FOR SEQ ID NO: 72:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 22 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
```

```

? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/457,648
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Petry, Douglas A.
? REGISTRATION NUMBER: 35,321
? REFERENCE/DOCKET NUMBER: 9205
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (510) 814-2974
? TELEFAX: (510) 814-2977
? INFORMATION FOR SEQ ID NO: 220:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 22 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-457-648-220

Query Match          90.0%; Score 18; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gcactgaagtaactaagg 20
DB 1 GCACGTGAAGTAACTAAGG 18

RESULT 9
? Sequence 72, Application US/08452055
? Patent No. 5705627
? GENERAL INFORMATION:
? APPLICANT: Bauer, Heidi M.
? APPLICANT: Greer, Catherine E.
? APPLICANT: Manos, Michele
? APPLICANT: Resnick, Robert M.
? APPLICANT: Ting, Yi
? TITLE OF INVENTION: Detection of Human Papillomavirus by the
? TITLE OF INVENTION: Polymerase Chain Reaction
? NUMBER OF SEQUENCES: 85
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hoffmann-La Roche Inc.
? STREET: 340 Kingsland Street
? CITY: Nutley
? STATE: New Jersey
? COUNTRY: U.S.A.
? ZIP: 07110
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/452,055
? FILING DATE:
? CLASSIFICATION: 536
? ATTORNEY/AGENT INFORMATION:
? NAME: Stias, Stacey R.
? REGISTRATION NUMBER: 32,630
? REFERENCE/DOCKET NUMBER: 9188
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (510) 814-2863
? TELEFAX: (510) 814-2977
? INFORMATION FOR SEQ ID NO: 72:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 22 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
```

MOLECULE TYPE: DNA (genomic)
US-08-452-055-72

Query Match 90.0%; Score 18; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gacactgaagtaactaagg 20
|||||
DB 1 GCACCTGAAGTAAGTAAGG 18

RESULT 10
US-08-545-809A-58
Sequence 58, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
TITLE OF INVENTION: HUMAN IMKUNOGLOBULIN VH GENE
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 629 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: human lymphoblast
CELL LINE: GM1
US-08-545-809A-58

Query Match 76.0%; Score 15.2; DB 3; Length 629;
Best Local Similarity 85.0%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 atgcactgaagtaactaagg 20
|||||
DB 53 ATGCACCTGAAGTAAGTAAGG 72

RESULT 11

US-08-433-126A-63
Sequence 63, Application US/08433126A
Patent No. 568935

GENERAL INFORMATION:
APPLICANT: STEPHENS, ANDREW
APPLICANT: SCHNEIDER, DAN
APPLICANT: GOLD, LARRY
TITLE OF INVENTION: NUCLEIC ACID LIGANDS OF TISSUE
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MG
COMPUTER: IBM pc compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,126A
FILING DATE: 03 MAY 1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX31.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 87 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
OTHER INFORMATION: All C's are 2'-F cytosine
FEATURE:
OTHER INFORMATION: All U's are 2'-F uracil
US-08-433-126A-63

Query Match 74.0%; Score 14.8; DB 1; Length 87;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 tgcactgaagtaactaagg 19
:|||||
DB 39 UGACUGAAGUACUAGG 56

RESULT 12
US-08-433-124A-63
Sequence 63, Application US/08433124A
Patent No. 5750342
GENERAL INFORMATION:
APPLICANT: STEPHENS, ANDREW
APPLICANT: SCHNEIDER, DAN
APPLICANT: GOLD, LARRY
TITLE OF INVENTION: NUCLEIC ACID LIGANDS OF TISSUE

```

? TITLE OF INVENTION: TARGET
? NUMBER OF SEQUENCES: 241
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Swanson & Bratschun, L.L.C.
? STREET: 8400 E. Prentice Avenue, Suite 200
? CITY: Englewood
? STATE: Colorado
? COUNTRY: USA
? ZIP: 80111
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MG
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: Wordperfect 6.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/433,124A
? FILING DATE: 03 MAY 1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/714,131
? FILING DATE: 10-JUNE-1991
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/536,428
? FILING DATE: 11-JUNE-1990
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/964,624
? FILING DATE: 21-OCTOBER-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Barry J. Swanson
? REGISTRATION NUMBER: 33,215
? REFERENCE/DOCKET NUMBER: NEX31.2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (303) 793-3333
? TELEFAX: (303) 793-3433
? INFORMATION FOR SEQ ID NO: 63:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 87 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? OTHER INFORMATION: All C's are 2'-F cytosine
? OTHER INFORMATION: All U's are 2'-F uracil
US-08-433-124A-63

? Query Match
? Best Local Similarity 74.0%; Score 14.8; DB 1; Length 87;
? Matches 12; Conservative 4; Mismatches 2; Indels 0; Caps 0;

QY 2 tgcactgaagtaactaag 19
? :||:||||:||||:|
Db 39 USGACUGAAGUACUAGG 56

RESULT 13
US-08-976-413A-63
? Sequence 63, Application US/08976413A
? Patent No. 6127119
? GENERAL INFORMATION:
? APPLICANT: STEPHENS, ANDREW
? APPLICANT: GOLD, LARRY
? TITLE OF INVENTION: NUCLEIC ACID LIGANDS OF TISSUE TARGET
? NUMBER OF SEQUENCES: 440
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Swanson & Bratschun, L.L.C.
? STREET: 8400 E. Prentice Avenue, Suite 200
? CITY: Englewood
? STATE: Colorado
? COUNTRY: USA
? ZIP: 80111
```

```

? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MG
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: Wordperfect 8.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/976,413A
? FILING DATE: 21-NOVEMBER-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/433,124
? FILING DATE: 03-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/714,131
? FILING DATE: 10-JUNE-1991
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/536,428
? FILING DATE: 11-JUNE-1990
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/964,624
? FILING DATE: 21-OCTOBER-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Barry J. Swanson
? REGISTRATION NUMBER: 33,215
? REFERENCE/DOCKET NUMBER: NEX31/CIP
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (303) 793-3333
? TELEFAX: (303) 793-3433
? INFORMATION FOR SEQ ID NO: 63:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 87 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? OTHER INFORMATION: All C's are 2'-F cytosine
? OTHER INFORMATION: All U's are 2'-F uracil
US-08-976-413A-63

? Query Match
? Best Local Similarity 74.0%; Score 14.8; DB 3; Length 87;
? Matches 12; Conservative 4; Mismatches 2; Indels 0; Caps 0;

QY 2 tgcactgaagtaactaag 19
? :||:||||:||||:|
Db 39 USGACUGAAGUACUAGG 56

RESULT 14
PCT-US96-06059-63
? Sequence 63, Application PC/TUS9606059
? GENERAL INFORMATION:
? APPLICANT: STEPHENS, ANDREW
? APPLICANT: SCHNEIDER, DAN
? TITLE OF INVENTION: NUCLEIC ACID LIGANDS OF TISSUE
? NUMBER OF SEQUENCES: 241
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Swanson & Bratschun, L.L.C.
? STREET: 8400 E. Prentice Avenue, Suite 200
? CITY: Englewood
? STATE: Colorado
? COUNTRY: USA
? ZIP: 80111
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MG
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: Wordperfect 6.0
? CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: PCT/US96/06059
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,124
; FILING DATE: 03-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,126
; FILING DATE: 03-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX31.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: All C's are 2'-F cytosine
; OTHER INFORMATION: All U's are 2'-F uracil
PCT-US96-06059-63

Query Match          74.0%; Score 14.8; DB 5; Length 87;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2   tgcactgaagtaactaag 19
          :|:|:|:|:|:|:|:|:|
Db      39   UGACUGAGACUACUAGG 56

RESULT 15
US-09-328-111-486/C
; Sequence 486, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steimann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroli III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 486
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(637)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-486

Query Match          74.0%; Score 14.8; DB 4; Length 637;
Best Local Similarity 88.9%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1   atgcactgaagtaactaa 18
          ||| ||| ||| ||| |||
Db      422   ATGCACCTGAGGAAACAAA 405

```

Search completed: April 3, 2002, 10:31:07
Job time: 3458 sec

Thu Apr 4 09:39:20 2002

us-09-885-799-317.rni

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 09:33:29 ; Search time 3285.41 Seconds
(Without alignments)
65.415 Million Cell updates/sec

Title: US-09-885-799-317

Sequence: 1 atgcacatgaagtaacaaag 20

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estl:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_pro:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	97.0	463	10	AM508191
2	17.4	89.0	547	10	AM508191
3	16.8	84.0	277	10	AM624074
4	16.8	84.0	307	10	AA324900
5	16.8	84.0	310	10	AA308882
6	16.8	84.0	324	11	T16859
7	16.8	84.0	326	11	H43832
8	16.8	84.0	334	10	AA908640
9	16.8	84.0	358	11	T58103
10	16.8	84.0	367	11	B1024053
11	16.8	84.0	369	11	BG008631
12	16.8	84.0	382	10	AM583616

13	16.8	84.0	388	11	BF202373
14	16.8	84.0	394	11	BG677300
15	16.8	84.0	396	10	AM583547
16	16.8	84.0	397	11	R50310
17	16.8	84.0	401	10	BF197720
18	16.8	84.0	406	10	AM376448
19	16.8	84.0	419	11	H43503
20	16.8	84.0	436	10	AA127111
21	16.8	84.0	440	10	AA448564
22	16.8	84.0	444	11	W73551
23	16.8	84.0	447	10	BE406937
24	16.8	84.0	448	11	W32860
25	16.8	84.0	458	10	AA448610
26	16.8	84.0	468	11	N73162
27	16.8	84.0	470	10	AA054781
28	16.8	84.0	470	10	AA150511
29	16.8	84.0	483	10	AA608747
30	16.8	84.0	493	10	AA043686
31	16.8	84.0	494	10	AA043686
32	16.8	84.0	498	10	BE429684
33	16.8	84.0	510	10	AA819870
34	16.8	84.0	510	11	N98715
35	16.8	84.0	527	10	AM829033
36	16.8	84.0	534	10	AM647909
37	16.8	84.0	586	13	A0019354
38	16.8	84.0	614	11	BG387057
39	16.8	84.0	618	10	AL037896
40	16.8	84.0	623	10	AW953082
41	16.8	84.0	637	10	AM604142
42	16.8	84.0	655	11	N98669
43	16.8	84.0	720	10	AU120061
44	16.8	84.0	722	11	BG328613
45	16.8	84.0	726	11	BF049955

ALIGNMENTS

RESULT 1
AM508191
LOCUS
DEFINITION
Gm-r1030-2338 5', mRNA sequence.
ACCESSION
AM508191
VERSION
AM508191.1 GI:7146206
KEYWORDS
SOURCE
ORGANISM
soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 463)

REFERENCE
AUTHORS
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepleow, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Ship, T., Jackson, Y., Cardenas, M., Mccann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com

Insert Length: 1960 Std Error: 0.00
High quality sequence stop: 454.

FEATURES

SOURCE

Location/Qualifiers
1. 463
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-T1030-2336"
/lab_host="DH10B"

BASE COUNT

195 a 67 c 113 g 88 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 18; DB 10; Length 463;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

2 tgcactgaagtaactaag 19

Db

380 TGCACGTGAGTAAGTAAAG 397

RESULT

2

LOCUS

AM624074 547 bp mRNA EST 18-MAY-2001

DEFINITION

EST322019 tomato flower buds 3-8 mm, Cornell University

ACCESSION

AM624074 Lycopersicon esculentum cDNA clone cTOB14D21 5', mRNA sequence.

VERSION

AM624074.1 GI:7337101

KEYWORDS

EST.

SOURCE

tomato.

ORGANISM

Lycopersicon esculentum

REFERENCE

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Asteridae: euasterids I; Solanales; Solanaceae; Solanum;

AUTHORS

van der Hoeven, R.S., Bezzeredes, J.L., Matern, A.L., Holt, I.E., Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Rohnig, C.M., Niernan, M., Fraser, C.M., Martin, G.B., Giovannoni, J.J., and Tanksley, S.D.

TITLE

Generation of ESTs from tomato flower tissue, 3-8 mm buds

JOURNAL

unpublished (1999)

COMMENT

Contact: CUCI

FEATURES

Location/Qualifiers

source

1. 547

/organism="Lycopersicon esculentum"

/db_xref="taxon:3847"

/db_xref="taxon:4081"

/clone="cTOB14D21"

/clone_lib="tomato flower buds 3-8 mm, Cornell University"

/dev_stage="3-8mm buds"

/note="Vector: pBluescript SK(-); site_1: EcoRI; site_2:

/note="Vector: pBluescript SK(-); site_1: EcoRI; site_2:

BASE COUNT

162 a 95 c 126 g 164 t

ORIGIN

Query Match

Best Local Similarity 94.7%; Score 17.4; DB 10; Length 547;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

2 tgcactgaagtaactaag 20

Db

520 TTCACGTGAGTAAGTAAAG 538

RESULT

3

LOCUS

AA347960 277 bp mRNA EST 21-APR-1997

DEFINITION

EST54368 Fetal heart II Homo sapiens cDNA 5' end, mRNA sequence.

ACCESSION

AA347960

VERSION

AA347960.1 GI:2000217

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS

1 (bases 1 to 277)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fudner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocke, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, W.-J., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Phillorino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bedarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, P., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Pannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.

TITLE

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL

Nature 377 (6547 Suppl.), 3-174 (1995)

MEDLINE

96026280

COMMENT

Other ESTs: THC171208

CONTACT: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tadb/hgi/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

source

1. 277

/organism="Homo sapiens"

/db_xref="ATCC (inost):149038"

/db_xref="taxon:9606"

/clone_lib="Fetal heart II"

/dev_stage="fetus"

/note="Organ: heart; Vector: pBluescript SK(-); site_1: EcoRI; site_2: XhoI"

BASE COUNT

64 a 59 c 67 g 85 t 2 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 277;
 Best Local Similarity 90.0%; Pred. No. 5.1e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgcactgaactgaactg 20
 |||||
 Db 92 ATGCACCTGAAGTACTCTGG 73

RESULT 4
 AA324900/c 307 bp mRNA EST 20-APR-1997
 LOCUS EST127802 Cerebellum II Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA324900
 VERSION AA324900.1 GI:1977144
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 307)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man'wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodex,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.F., Fertle,A., Fischer,C., Hastings,G.A., He,M.W.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

JOURNAL
 MEDLINE
 COMMENT 96026280
 Other ESTs: FHC171208
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES
 source
 Location/Qualifiers
 1..307
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):125393"
 /db_xref="taxon:9606"
 /clone_lib="Cerebellum II"
 /tissue_type="Cerebellum"
 /dev_stage="adult"
 /note="Organ: brain; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"

BASE COUNT 69 a 67 c 77 g 91 t 3 others
 ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 307;
 Best Local Similarity 90.0%; Pred. No. 5.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgcactgaactgaactg 20
 |||||
 Db 34 ATGCACCTGAAGTACTCTGG 15

RESULT 5
 AA308882/c 310 bp mRNA EST 18-APR-1997
 LOCUS EST179686 Cerebellum II Homo sapiens cDNA 5' end similar to EST
 DEFINITION containing Alu repeat, mRNA sequence.
 ACCESSION AA308882
 VERSION AA308882.1 GI:1961220
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 310)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man'wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodex,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.F., Fertle,A., Fischer,C., Hastings,G.A., He,M.W.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kun,L.C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

JOURNAL
 MEDLINE
 COMMENT 96026280
 Other ESTs: FHC171208
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES
 source
 Location/Qualifiers
 1..310
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):127113"
 /db_xref="taxon:9606"
 /clone_lib="Cerebellum II"
 /tissue_type="Cerebellum"
 /dev_stage="adult"
 /note="Organ: brain; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"

BASE COUNT 66 a 66 c 81 g 97 t
 ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 310;
 Best Local Similarity 90.0%; Pred. No. 5.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgcactgaactgaactg 20
 |||||
 Db 146 ATGCACCTGAAGTACTCTGG 127

TITLE The Masnu-Merck EST Project
JOURNML Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
High quality sequence stops: 318
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: M13p1
High quality sequence stop: 318.
Location/Qualifiers
1..326
/organism="Homo sapiens"
/db_xref="db:3819001"
/db_xref="taxon:9606"
/clone="IMAGE:188104"
/clone.lib="Soares breast 3NBHBL"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pUT73D (Pharmacia) with a
modified polylinker; Site:1: Not 1; Site:2: Eco RI; 1st
strand cDNA was primed with a Not 1 - oligo(dT) primer [5
TCCTACCACTGACAGCTGAGCGCCGCCCTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not 1 and cloned into the Not 1
and Eco RI sites of a modified pUT73 vector (Pharmacia).
Library went through one round of normalization to a Cot
20. Library constructed by Bento Soares and M.Fatima
Ronalds."

BASE COUNT 68 a 71 c 76 g 99 t 12 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 11; Length 326;
Best Local Similarity 90.0%; Pred No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

OY 1 atgcaactgaagtaactaagg 20
|||||
Db 199 ATGCACTGAGAGTAAGTCTGCG 180

RESULT 8
AA908640/c 334 bp mRNA EST 09-JUN-1998
LOCUS 0104a02.s1 NCI-CCAP-Lu5 Homo sapiens cDNA clone IMAGE:1522442 3',
DEFINITION
RNA sequence.
ACCESSION AA908640
VERSION
KEYWORDS AA908640.1 GI:3048045
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 334)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CCAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strussberg, Ph.D.
Email: ccaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I M A C P Consortium at <http://www.imacp.org>

www.bio.livnl.gov/bdrrp/image/image.html
 Insert Length: 1020 Std Error: 0.00
 Seq primer: 40m13 fwd. ET from Amersham
 High quality sequence stop: 305.
 Location/Qualifiers

FEATURES

1..334

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:1522442"
 /clone_1ib="NCI.CGAP.Lu5"
 /issue_type="carcinoid"
 /lab_host="DH10B"

/note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - Oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

76 a 75 c 81 g 102 t

Query Match 84.0%; Score 16.8; DB 10; Length 334;
 Best Local Similarity 90.0%; Pred. No. 5.3e+02;
 Matches 18: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgcacggaagtaactaagg 20
 ||||||||||||||||
 Db 67 ATGCACGTGAGTACTCTCG 48

RESULT

9

LOCUS

TS8103/c

DEFINITION

YB26C03.F1 Striatagene fetal spleen (#937205) Homo sapiens cDNA

clone IMAGE:72292 5' similar to contains LTR6 repetitive element,

MRNA sequence.

ACCESSION

TS8103

VERSION

TS8103.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 358)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapelli, B.,

Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins

, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Maritz, E., Moore

, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,

Scheibenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,

Underwood, K., Womlamann, P., Waterston, R., Wilson, R. and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

CONTACT: Wilson RK

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

High quality sequence stops: 311

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.livnl.gov) for further information.

FEATURES

1..358

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="GDB:43957"

/db_xref="taxon:9606"

BASE COUNT

66 a 98 c 110 g 93 t

/clone="IMAGE:72292"
 /clone_1ib="Striatagene fetal spleen (#937205)"
 /issue_type="fetal spleen"
 /dev_stage="fetal"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: spleen; Vector: pBluescript SK-; Site: 1;
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. Pooled spleens. Average insert size: 1.0 kb;
 Uni-ZAP XR Vector; -5' adaptor sequence: 5' GATATCGGACGAG
 3' -3' adaptor sequence: 5' CTCGACATTTTTTTTTTTTTTTT 3'."

BASE COUNT

68 a 87 c 88 g 115 t

Query Match 84.0%; Score 16.8; DB 11; Length 358;
 Best Local Similarity 90.0%; Pred. No. 5.3e+02;
 Matches 18: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgcacggaagtaactaagg 20
 ||||||||||||||||
 Db 310 ATGCACGTGAGTACTCTCG 291

RESULT

10

LOCUS

B1024053/c

DEFINITION

CMO-MT0353-240101-755-cl1 MT0353 Homo sapiens cDNA, mRNA sequence.

ACCESSION

B1024053

VERSION

B1024053.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 367)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CMO&t2=CMO-MT0353-

240101-755-cl1&t3=2001-01-24&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 367.

FEATURES

1..367

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_1ib="MT0353"

/dev_stage="Adult"

/note="Organ: marrow; Vector: puc18; Site: 1; SmaI; Site: 2;

SmaI; A mini-library was made by cloning products derived

from ORFESTES PCR (U.S. Letters Patent application No. 196

,716 - Ludwig Institute for Cancer Research) profiles

into the pUC 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 388)
AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsja, C.C., Kang, Y., Lazo, G.R., Miller, R., Kausch, C.J., Seaton, C.L. and Tong, J.C.
TITLE The structure and function of the expressed portion of the wheat genomes - pre-anthesis spike cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@w.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES
source
1..388
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHEI767_H07_013"
/tissue_lib="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Whole spike with awns trimmed, while, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the T7 Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
BASE COUNT 92 a 94 c 111 g 89 t 2 others
ORIGIN
Query Match 84.0%; Score 16.8; DB 11; Length 388;
Best Local Similarity 90.0%; Pred. No. 5,4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
YQ 1 atgacatgaactaactaagg 20
|||||
Db 344 ATGACGCTGAGGATCTAAGC 363
RESULT 14
Bg677300 394 bp mRNA EST 01-MAY-2001
LOCUS 602624012F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4748979 5',
DEFINITION mRNA sequence.
ACCESSION Bg677300
VERSION Bg677300.1 GI:13908697
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 394)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov

Tissue Procurement: James Cleaver, M.D.
cDNA library preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LAM10601 row: h column: 04
High quality sequence stop: 393.
FEATURES
source
1..394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4748979"
/tissue_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Skin; Vector: PCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 83 a 93 c 96 g 122 t
ORIGIN
Query Match 84.0%; Score 16.8; DB 11; Length 394;
Best Local Similarity 90.0%; Pred. No. 5,4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
YQ 1 atgacatgaactaactaagg 20
|||||
Db 312 ATGACGCTGAGGATCTGCG 293
RESULT 15
AM583547 396 bp mRNA EST 05-JUL-2001
LOCUS ia01a04.y1 Human Pancreatic Islets Homo sapiens cDNA 5', mRNA
DEFINITION sequence.
ACCESSION AM583547
VERSION AM583547.1 GI:7260505
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 396)
AUTHORS Melton, D., Brown, J., Kently, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, L., Searce, M., Brestelli, J., Gradow, G., Clifton, S., Hillier, L., Maira, M., Pape, D., Wyle, T., Martin, J., Bistacchi, A., Schmitt, A., Theising, B., Riller, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagariswilli, R., Williams, T., Jackson, Y. and Bowers, Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biochem.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown (brownj@fas.harvard.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 395.
FEATURES
source
1..396
/organism="Homo sapiens"
/db_xref="taxon:9606"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 11:00:04 ; Search time 1711.61 Seconds

(without alignments)
192.768 Million cell updates/sec

Title: US-09-885-799-488

Sequence: 20

1 gaacggcattactctctgt 20

Scoring table:

IDENTITY_NUC

Searched:

1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pal:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_sy:*

28: em_un:*

29: em_vl:*

30: em_hgtgo_hum:*

31: em_hgtgo_inv:*

32: em_hgtgo_rod:*

33: em_hgtgo_hum:*

34: em_hgtgo_inv:*

35: em_hgtgo_rod:*

36: em_hgtgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	415	14	HP01535	U01535 Human papillomavirus
2	20	100.0	455	14	HP012476	U12476 Human papillomavirus
3	20	100.0	455	14	HP0124860	U12486 Human papillomavirus
4	20	100.0	1059	6	AR084349	AR084349 Sequence
5	20	100.0	1521	6	AR084348	AR084348 Sequence
6	20	100.0	3283	6	AR084352	AR084352 Sequence
7	20	100.0	3283	14	HP022461	U22461 Human papillomavirus
8	20	100.0	7905	14	HP021941	U21941 Human papillomavirus
9	16.8	84.0	394	6	I22791	I22791 Sequence 27
10	16.8	84.0	394	6	I47616	I47616 Sequence 27
11	16.8	84.0	2535	10	MMU011024	AJ124830 Mus musculus
12	16.8	84.0	4953	10	MMU242830	AJ124830 Mus musculus
13	16.8	84.0	35000	1	AB000617	AB000617 Bacillus
14	16.8	84.0	63731	2	AL360016_4	Continuation (5 of
15	16.8	84.0	146300	9	AL157903	AL157903 Human DNA
16	16.8	84.0	171415	2	AP002771	AP002771 Homo sapi
17	16.8	84.0	193390	2	AP002771	AP002771 Homo sapi
18	16.8	84.0	206647	9	AP002898	AP002898 Homo sapi
19	16.8	84.0	221160	1	BSUB00002	Z99105 Bacillus su
20	16.8	84.0	338100	1	TACID2	AL445064 Thermopila
21	16.4	82.0	993	1	STRATPCA	L08890 Salmonella
22	16.4	82.0	6601	1	ARGFCRABC	M93187 Arthrobacte
23	16.4	82.0	16832	1	AF042490	AF042490 Arthrobac
24	16.4	82.0	105964	9	AL135940	AL135940 Human DNA
25	16.4	82.0	190839	2	AL365402	AL365402 Homo sapi
26	16.4	82.0	216431	2	AL354771	AL354771 Homo sapi
27	16	80.0	4582	10	MUSPLM24X	M98502 Mus musculus
28	16	80.0	227074	2	AC060772	AC060772 Mus muscu
29	15.8	79.0	859	4	SSCASEIN	X51977 Pig RNA fo
30	15.8	79.0	2134	8	AF132121	AF132121 Pinus lae
31	15.8	79.0	8471	5	IPIGHD	X52617 Catfish lph
32	15.8	79.0	10372	1	AE003953	AE003953 xyella f
33	15.8	79.0	12575	5	LEPGMGHC	X79482 I-punctatus
34	15.8	79.0	14222	1	AE004401	AE004401 Vibrio ch
35	15.8	79.0	40744	1	SC04	AL109590 Streptomy
36	15.8	79.0	68125	2	AC020345	AC020345 Streptophi
37	15.8	79.0	105866	2	AC021602	AC021602 Homo sapi
38	15.8	79.0	106742	9	AC008904	AC008904 Homo sapi
39	15.8	79.0	106750	2	AC008943	AC008943 Homo sapi
40	15.8	79.0	143290	2	AP003825	AP003825 Oryza sat
41	15.8	79.0	148540	9	HS212P9	AL009181 Human DNA
42	15.8	79.0	149718	2	AC016633	AC016633 Homo sapi
43	15.8	79.0	154586	2	AC016634	AC016634 Homo sapi
44	15.8	79.0	156586	2	AC026350	AC026350 Homo sapi
45	15.8	79.0	167335	2	AC023527	AC023527 Homo sapi

ALIGNMENTS

RESULT 1

LOCUS HP01535 415 bp DNA

DEFINITION Human papillomavirus PCR type A01 sequence.

ACCESSION U01535

VERSION U01535.1 GI:437112

KEYWORDS

SOURCE

ORGANISM Human papillomavirus.

VIRUSES: dsDNA viruses, no RNA stage; Papillomaviridae;

Papillomavirus.

REFERENCE 1 (bases 1 to 415)

AUTHORS Tachezy, R., Van Ranst, M.A., Cruz, Y. and Burk, R.D.

TITLE Analysis of short novel human papillomavirus sequences

JOURNAL Biochem. Biophys. Res. Commun. 204 (2), 820-827 (1994)

MEDLINE 95071400

REFERENCE 2 (bases 1 to 415)

AUTHORS Van Ranst, M.A.

TITLE Direct Submission

JOURNAL Submitted (07-SEP-1993) Marc A. Van Ranst, Dept. of Microbiology &

Immunology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA

FEATURES

source

1. .415
/organism="Human papillomavirus"
/specific host="human"
/db_xref="taxon:10566"
/note="HPV PCR type A21 as amplified with MY09 and MY11 primers"

BASE COUNT 132 a 78 c 78 g 127 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 415;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaacggcattaccctgctgt 20
|||||
Db 80 GAAACGGCATTACTGCTGT 99

RESULT 2

HPV12476 455 bp DNA VRL 24-MAY-1995
LOCUS Human papillomavirus isolate Cp141 L1 protein gene, MY09/MY11
DEFINITION region, partial cds.
ACCESSION U12476
VERSION U12476.1 GI:577354
KEYWORDS Human papillomavirus.
SOURCE Human papillomavirus.
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.

REFERENCE 1 (bases 1 to 455)
AUTHORS Peyton, C.L. and Wheeler, C.M.
TITLE Identification of five novel human papillomavirus sequences in the New Mexico triethnic population
JOURNAL J Infect. Dis. 170 (5), 1089-1092 (1994)

MEDLINE 95052823
REFERENCE 2 (bases 1 to 455)
AUTHORS Farmer, A.D.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-1994) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA

FEATURES
source Location/Qualifiers
1. .455
/organism="Human papillomavirus"
/proviral
/isolate="Cp141"
/db_xref="taxon:10566"
<1. .>455
/codon_start=1
/product="L1 protein, MY09/MY11 region"
/protein_id="AA67220.1"
/db_xref="GI:577355"
/translation="AAGTNGICMHNQLFITVDTTRSTNFTLSACTETALPAVVSPT
KFEKVRHVEEDLOFLFOICTITLADVMAYIHMTNPAIIDMNNGVPPPSASLVD
TYRLOSAAIVCOKDAPTPEKRDYDLPKFNVDLKEKFTSLDQPLGR"

CDS

BASE COUNT 142 a 86 c 90 g 137 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaacggcattaccctgctgt 20
|||||
Db 100 GAAACGGCATTACTGCTGT 119

RESULT 3

HPV124860 455 bp DNA VRL 24-MAY-1995
LOCUS Human papillomavirus isolate LVX160 L1 protein gene, MY09/MY11
DEFINITION region, partial cds.
ACCESSION U12486
VERSION U12486.1 GI:577374
KEYWORDS Human papillomavirus.
SOURCE Human papillomavirus.
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.

REFERENCE 1 (bases 1 to 455)
AUTHORS Ong, C.K., Bernard, H.U. and Villa, L.L.
TITLE Identification of genomic sequences of three novel human papillomavirus sequences in cervical smears of Amazonian Indians
J. Infect. Dis. 170 (5), 1086-1088 (1994)

MEDLINE 95052822
REFERENCE 2 (bases 1 to 455)
AUTHORS Farmer, A.D.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-1994) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA

COMMENT This entry was obtained from cervical smears taken from members of isolated Amazonian tribes. The samples were PCR-amplified using the MY09/MY11 consensus primers, then examined in hybridization experiments in order to determine their homology with known HPV isolates. This novel variant is more than 10% divergent from its closest known HPV variant, suggesting that it qualifies as a new isolate. Although the tribes were thought to have been sexually isolated from non-American Indian populations for at least 12,000 years, sequences closely related to these novel variants have since been detected in other distinct populations. This may be evidence for the hypothesis that papillomavirus isolates evolved before the speciation of Homo sapiens, and consequently before the divergence of ethnic groups.

FEATURES
source Location/Qualifiers
1. .455
/organism="Human papillomavirus"
/proviral
/isolate="LVX160"
/db_xref="taxon:10566"
/tissue_type="cervix"
<1. .>455
/codon_start=1
/product="L1 protein, MY09/MY11 region"
/protein_id="AA67230.1"
/db_xref="GI:577375"
/translation="AAGTNGICMHNQLFITVDTTRSTNFTLSACTETALPAVVSPT
KFEKVRHVEEDLOFLFOICTITLADVMAYIHMTNPAIIDMNNGVPPPSASLVD
TYRLOSAAIVCOKDAPTPEKRDYDLPKFNVDLKEKFTSLDQPLGR"

CDS

BASE COUNT 142 a 86 c 90 g 137 t
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaacggcattaccctgctgt 20
|||||
Db 100 GAAACGGCATTACTGCTGT 119

RESULT 4

AR084349 1059 bp DNA PAT 01-SEP-2000
LOCUS Sequence 11 from patent US 5981173.
DEFINITION AR084349
ACCESSION AR084349
VERSION AR084349.1 GI:1001120
KEYWORDS Unknown.
SOURCE

```

ORGANISM      Unknown.
REFERENCE     1 (bases 1 to 1059)
AUTHORS       Orth,G., Beaudenon,S. and Longuet,M.
TITLE         Genital human papillomavirus type 68a (HPV-68a), related to the
              potentially oncogenic HPV-39
JOURNAL       Patent: US 5981173-A 11 09-NOV-1999;
FEATURES      Location/Qualifiers
SOURCE        1..1059
              /organism="unknown"
BASE COUNT    291 a 207 c 244 g 317 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 1059;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gaaacggcatatcctctgt 20
Db 782 GAAACGCCATACCTGCTGT 801

RESULT 5
AR084348      1521 bp      DNA
LOCUS         Sequence 10 from patent US 5981173.
DEFINITION    AR084348
ACCESSION     AR084348.1 GI:10011119
VERSION
KEYWORDS
SOURCE        Unknown.
ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 1521)
AUTHORS       Orth,G., Beaudenon,S. and Longuet,M.
TITLE         Genital human papillomavirus type 68a (HPV-68a), related to the
              potentially oncogenic HPV-39
JOURNAL       Patent: US 5981173-A 10 09-NOV-1999;
FEATURES      Location/Qualifiers
SOURCE        1..1521
              /organism="unknown"
BASE COUNT    419 a 304 c 348 g 450 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 1521;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gaaacggcatatcctctgt 20
Db 1045 GAAACGCCATACCTGCTGT 1064

RESULT 6
AR084352      3283 bp      DNA
LOCUS         Sequence 14 from patent US 5981173.
DEFINITION    AR084352
ACCESSION     AR084352
VERSION
KEYWORDS
SOURCE        Unknown.
ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 3283)
AUTHORS       Orth,G., Beaudenon,S. and Longuet,M.
TITLE         Genital human papillomavirus type 68a (HPV-68a), related to the
              potentially oncogenic HPV-39
JOURNAL       Patent: US 5981173-A 14 09-NOV-1999;
FEATURES      Location/Qualifiers
SOURCE        1..3283
              /organism="unknown"
BASE COUNT    945 a 640 c 699 g 999 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 3283;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gaaacggcatatcctctgt 20
Db 1070 GAAACGCCATACCTGCTGT 1089

RESULT 7
HPU22461      3283 bp      DNA
LOCUS         Human papillomavirus type 70 E6 protein (E6), E7 protein (E7), and
DEFINITION    L1 protein (L1) genes, complete cds.
ACCESSION     U22461
VERSION       U22461.1 GI:1345087
KEYWORDS
SOURCE        Human papillomavirus type 70.
              Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
              Papillomavirus.
REFERENCE     1 (bases 1 to 3283)
AUTHORS       Longuet,M., Beaudenon,S. and Orth,G.
TITLE         Two novel genital human papillomavirus (HPV) types, HPV68 and
              HPV70, related to the potentially oncogenic HPV39
JOURNAL       J. Clin. Microbiol. 34 (3), 738-744 (1996)
MEDLINE       97060129
AUTHORS       Longuet,M.
REFERENCE     2 (bases 1 to 3283)
              Direct Submission
              Submitted (13-MAR-1995) Michele Longuet, Laboratoire des
              Papillomavirus, Unite INSERM 190, Institut Pasteur, 25 rue du Dr
              Roux, Paris 75015, France
JOURNAL
FEATURES      Location/Qualifiers
SOURCE        1..3283
              /organism="Human papillomavirus type 70"
              /db_xref="taxon:39457"
              /gene="L1"
              /gene="L1"
              /gene="L1"
              /note="L1"
              /note="putative major capsid protein"
              /codon_start=1
              /product="L1 protein"
              /protein_id="AAC54879.1"
              /db_xref="GI:1345088"
              /translation="MALRRSSDNTVYLRPPSVAKVNTDDYVTRGTYVYVAGTRELTL
              VGHPFVYPVNGRKOELPKVSAYRVFRVSLPDPNKEQLPPSLYNDPQRLVWMC
              IGEVIGRGOPLGVSAGHPLVNRLLDTFNSHFSVSVYQUSRNVSVMYKOTGLCTIG
              CVPAMGEGMAKAKKSTTVQOQDCPLELNTALEDGDMDTGAGMDPRLDQETS
              EVDLDIGOSCKYPDYLQMSADYVDSMFLLRKEQLFAHFMNNGCAVGTISELX
              IKGTDIREPGTHVSPSSGMSVSDSLFNKRYLRLHKAQGNNGICMHNQLFTIV
              DTRSTNFTLSACTETAIIPAVYSPYKEEYTRHVEYDLOFIPDLCTITLADAVAVI
              HTMPALIDMNGIVTPPSASLVDTYRYQSAIAQCKAPPEKQVYDILKENVY
              DLKRFSTELDQPIAKRFLAQVGARRRPTIGPKRPAKSSSSASAKHKKRYSK"
              2419..2425
              /note="putative"
              2454..2930
              /gene="E6"
              2454..2930
              /gene="E6"
              /note="putative transforming protein"
              /codon_start=1
              /product="E6 protein"
              /protein_id="AAC54880.1"
              /db_xref="GI:1345089"
              /translation="MARPPNPAERPYKLPDLCTALDITJHDIITIDCVCKTQLOOTEV
              YEFASDLPIVNGCEPYAACOKCIKFAKVRRLRHSNSVYATLTSESTNTKLYDS
              IRCMSCLKPAPKLRHVNTKRRFIQIQAISYIGQCRHCTSNREDRRIRRETV"
              2939..3268
              /gene

```


CDS	/gene="E7" 2939..3268 /note="putative transforming protein" /codon_start=1 /product="E7 protein" /protein_id="AAC54851.1" /db_xref="GI:1345090" /translation="MHGPRPTLOEIVLDLYPNEIQPVYDVCHROIEDSNENETDEPDR VYHQQOLARREPPQRHRIQCMCKCNTTLHLVVASQENIRSLQLEWETLSFVCP WASSTQ"
BASE COUNT	945 a 640 c 699 g 999 t
ORIGIN	WASSTQ
Query Match	100.0%; Score 20; DB 14; Length 3283; Best Local Similarity 100.0%; Pred. No. 2.4;
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OR	1 qaacggccataccctcgt 20
Db	1070 GAACGCGCATACCTCTGT 1089
RESULT	8
LOCUS	HP021941 7905 bp DNA VRI 30-MAY-1996
DEFINITION	Human papillomavirus type 70, complete genome.
ACCESSION	U21941
VERSION	U21941.1 GI:1173493
KEYWORDS	.
SOURCE	Human papillomavirus type 70.
ORGANISM	Human papillomavirus type 70. Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.
REFERENCE	1 (bases 1 to 7905) Forstlund, O. and Hansson, B.G. Human papillomavirus type 70 genome cloned from overlapping PCR products: complete nucleotide sequence and genomic organization J. Clin. Microbiol. 34 (4), 802-809 (1996)
JOURNAL	96249586
MEDLINE	2 (bases 1 to 7905) Forstlund, O.
REFERENCE	Direct Submission Submitted (01-MAR-1995) Ola Forstlund, Clinical Virology, Medical Microbiology, Malmo University Hospital, Ing. 78 MAS, Malmoec, S-205 02, Sweden On Feb 2, 1996 this sequence version replaced gi:1160342. Location/Qualifiers
FEATURES	1..7905 /organism="Human papillomavirus type 70" /db_xref="taxon:394577" /note="Isolated from a cervical brush sample from a patient with diagnosis of cervical condyloma; from cloned overlapping PCR products" 107..583 /gene="E6" 107..583 /gene="E6" /note="Method: conceptual translation supplied by author.; putative E6" /codon_start=1 /protein_id="AAC54850.1" /db_xref="GI:717152" /translation="MARPNARPRPYLPPLCALHTLHDTLIDCVCKTQLOOTVY YERASDELIVRKEPFAACQCKIKFHAVERELRHSNSVATTLLESITVTKLYNIS IRKMSLKLPCAEKLELRHVNTKRREHQLAGSYGQCRCWTSNRRRRIRRETVQ" 592..921 /gene="E7" 592..921 /gene="E7" /note="Method: conceptual translation supplied by author.; putative E7"
gene	
CDS	
gene	/codon_start=1 /protein_id="AAC54851.1" /db_xref="GI:717153" /translation="MHGPRPTLOEIVLDLYPNEIQPVYDVCHROIEDSNENETDEPDR VYHQQOLARREPPQRHRIQCMCKCNTTLHLVVASQENIRSLQLEWETLSFVCP WASSTQ" join(928..2746,2748..2886) /gene="E1" 928..2886 /gene="E1" /note="Method: conceptual translation supplied by author.; putative E1" /codon_start=1 /protein_id="AAC54852.1" /db_xref="GI:1160343" /translation="MANCEGTGDDSGCGNGEWELVQALIVDKOTGDTVSEDEENATDNG SDLAFLDPTDTCVOAERERAVQVLYNMOEADRDASVALRKRYGCSNIAKSCQAP PGVHREGRYTLOELPYNICCNQARTNYSVPLSGYGNMEVTADEVTVVNTNGEBS" GENGENGCSIRECCSSVDSADISBENODOSPTRLKTVLQANNAKALLISQFKNYGG LARDLVRTRFSKDTICTDWAAATQVNAITAGETLTLQPYALTLHQCLDITGYGVY ITLLIRKCGKNITVGGKLSKLHVPSCMLTEPRILASPYALAYRTQMSNISEV SGTPPEIQRLLVYIQHGDIDSVFPLSDMVQMAFNDVDESDIAYGALADUSNSNA APLKSNOCAKAYVDCATCMRHYKRAQKQKQMTMAQWTRFPCDQCDGCDMPRIYQFLY QGYEFTFLCAKEKEPLKGPCKKNCIVIOGPPNTGSCYFCMSLMHFOCTGVSYNYS HEMLLEPLADAKVAMLDADATGTCWSYFDTWRNLDGNPISIDRKRIHILQIKPPLIT TSNTNPEERBRPYLTSRLVYFTFPNAPFPDDQKNRYIYIINKMKSFQATWCKIDL. QQDEDEGDDNGNTYLPFCVGTGENTRTL" join(2813..3079,3081..3260,3262..3895) /gene="E2" 2813..3895 /gene="E2" /note="Method: conceptual translation supplied by author.; putative E2" /codon_start=1 /protein_id="AAC54853.1" /db_xref="GI:1160344" /translation="MKETMETLSQRLNALQOEILFHYEODSKLTYDQINWKYVRL NAIFVAREGKMTIDHOVYPPGTTSKAKAYQALIELOMALESLATOPNKEEMTKCT SNEMOTIKPKOCEKKGGYVAVWDKKNMSMYVWGAIVYKKTHTPWIKKTRSYNDY WGIYVYHQBKTYVYVFKQDQNGYTGSKRREYCNENIICHDPSVTSIDVYPTTHL TALQHTTPARTAAATTPCTKTKASAPSCQGVSRPSETVGVVDLVTSKGNRRKQQ GSDTPTVHLAKGDKGKGLCLRLKRNLSYENISCTMWHIGKSKHTGIIIVYVYT EAQRQKFLTVRIPPSVHVSVMTL" 3408..3662 /gene="E3" 3408..3662 /gene="E3" /note="Start codon absent. Method: conceptual translation supplied by author.; putative E3" /codon_start=1 /protein_id="AAC54854.1" /db_xref="GI:717156" /translation="FIVITLCTQVVTTOYPLLSILQNTPEPRPEIPQVHAAKRIKSR RRLASVSPDPKQKOTECSWTLLQVKAATNDGSIYVYTHL" 3909..4145 /gene="E5" 3909..4145 /gene="E5" /note="Method: conceptual translation supplied by author.; putative E5" /codon_start=1 /protein_id="AAC54855.1" /db_xref="GI:717157" /translation="MYIVYISMALVFLWFAVCIYICGVSPLPSVHLCAVWMLILP VEIVVHTPPLQMPICVLYLFLPMWFLHLSVYA" join(4209..4752,4754..5174,5176..5560,5562..5609) /gene="I2" 4209..5609 /gene="I2" /note="Method: conceptual translation supplied by author.; putative I2" /codon_start=1 /protein_id="AAC54856.1"
CDS	
gene	
CDS	

/db_xref="GI:1173494"
/translation="WSSPASRRKASATIDYTKCKGCPEDVYVNEGTTLADE
LQMSLGIETLGIIGTGTGGRGTYLPIAGPSTVTVTPARPVTEVCTERS
IVOLVESVSSGTPPTPTGTSGPETSATITPAVDITPASGVQISTSTNP
AFADPSLIEVPOTGEVSGNITVTPTGTGTELEIMQVFAHGTEPISSTVNP
SRVAPRLYSRAHVRVNNEDVTPSPFVDNAPSPGDTSLTPPADTADDPF
LDIVLHRPALTISRQVFRSLKATMTFRTGTGAVHYHDIENITATEDIM
OPLTSESDGLYDIADADIDNMLHTSTGTSRSHSPSPSTVSTKSNIT
IPFTSMIDIPVTGPDILVPLPASPNIPLFPPTSIDTTVAIAGSNITLLPLTYFLK
KKRRVPTDGPVAV"
join(5590..6784,6786..7104)
/gene="L1"
5590..7104
/gene="L1"
/note="Method: conceptual translation supplied by author.;
putative L1"
/codon_start=1
/protein_id="AAC54857.1"
/db_xref="GI:1160346"
/translation="MALWSSDNTVLPPEPVAKVYNTDVTHTGIVYVAGSSRLIT
VGHPEKVPVNGRKQELPKYSAYQVYRVSLLPPNKGFLPPSLNPDQLWVAC
IGYEIRGQPLCYVSGHPLNRLDITENSHSSAVSTQSDNVSVYKQTLGIG
CVPAMEHMAKGRACKSTQGDICPLEVNTALEGDMIDTGYGAMDPFRLQETKEV
PLDICQSVCKEPDYLOMSADYGDSPFLKLEQLFAHPNMGAMGDTLPSELYK
GTDIREPCTHYVSPSSKSVSSDQLFKFMYLHAKOHNHNGMNDPLTIVYT
TRSTNFTLSACTETAIIPAVYSPTEKREYRHYEEDLQIFLOLCTITLTDVMAVYHT
MNPAILDMNNGVTPPSASLVPTRYRLOSALACQKDPPEKDPDLDLKWVNDL
KEKSTELDQFLGKRFLLQVGARRRTIGPKRRPASKSSSSSKRRKRVSK"

gene
CDS

BASE COUNT 2425 a 1495 c 1693 g 2292 t
ORIGIN

Query Match 100.0% Score 20; DB 14; Length 7905;
Best Local Similarity 100.0% Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaacagccatacctcgtgt 20
|||||
Db 6628 GAAACGCCACTGCTGT 6647

RESULT 9
LOCUS 122791 394 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 279 from patent US 5527898.
ACCESSION 122791
VERSION 122791.1 GI:1603145
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 394)
AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Manos,M.Michele, Kesnick,R.M.
and Zhang,T.Y.
TITLE Detection of human papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5527898-A 279 18-JUN-1996;
FEATURES
source 1..394
location/Qualifiers
BASE COUNT 127 a 80 c 74 g 113 t
ORIGIN

Query Match 84.0% Score 16.8; DB 6; Length 394;
Best Local Similarity 90.0% Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaacagccatacctcgtgt 20
|||||
Db 41 GAAACGCCACTGCTGT 60

RESULT 10

147616
LOCUS 147616 394 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 279 from patent US 5639871.
ACCESSION 147616
VERSION 147616
KEYWORDS 147616.1 GI:2471581
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 394)
AUTHORS Bauer,H.M., Gravitt,P.E., Greer,C.E., Imprim,C.C.,
Manos,M.Michele, Kesnick,R.M., and Zhang,T.Y.
TITLE Detection of human papillomavirus by the polymerase chain reaction
JOURNAL Patent: US 5639871-A 279 17-JUN-1997;
FEATURES
source 1..394
location/Qualifiers
BASE COUNT 127 a 80 c 74 g 113 t
ORIGIN

Query Match 84.0% Score 16.8; DB 6; Length 394;
Best Local Similarity 90.0% Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaacagccatacctcgtgt 20
|||||
Db 41 GAAACGCCACTGCTGT 60

RESULT 11
LOCUS MM0011024 2535 bp DNA ROD 02-JUN-1999
DEFINITION Mus musculus Ert-1 and Ert-2 gene transcripts, exons 1 to 2.
ACCESSION AJ011024
VERSION AJ011024.1 GI:4902855
KEYWORDS Ert-1 gene; Ert-2 gene; erythroid cell-specific and testis-specific
protein 1; erythroid cell-specific and testis-specific protein 2.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2535)
AUTHORS Xue,H., O'Neill,D., Morrow,J., and Rank,A.
TITLE A novel mouse gene, Hemi, encoding an hematopoietic cell-specific
transcript
JOURNAL Gene 231 (1-2), 49-58 (1999)
MEDLINE 99250249
REFERENCE 2 (bases 1 to 2535)
AUTHORS Xue,H.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-1998) Xue H., Genetics and Development, Columbia
University, Room 1614, HNSC, 701 West 168th Street, New York, New
York 10032, U.S.A
COMMENT Related sequences AJ011022 and AJ011023.
FEATURES
source 1..2535
location/Qualifiers
1..2535
/organism="Mus musculus"
/db_xref="taxon:10090"
753..1035
/gene="Ert-1"
/number=1
/number=1
753..998
/gene="Ert-2"
/number=1
join(753..998,999..1685,1686..2128)
/gene="Ert-2"
/note="Ert-1"
join(753..1035,1036..1685,1686..2128)
/gene="Ert-1"
join(753..1035,1686..2128)
/gene="Ert-1"
/product="erythroid cell-specific and testis-specific

MRNA	protein 1 (ERT-1)" join(753, .998,1686, .2128) /gene="Ert-2" /product="erythroid cell-specific and testis-specific protein 2 (ERT-2)" 837, .1034 /gene="Ert-1" /note="erythroid cell-specific and testis-specific gene transcript 1" /codon_start=1 /product="erythroid cell-specific and testis-specific protein 1 (ERT-1)" /protein_id="CA843517.1" /db_xref="GI:4902856" /translation="MPSRRGTSSNSRRSRNRNRRNSSSSCNSSNRNNSKKNKAKSN LQPGSGVGVCHSALICQGLHQGLHYL" join(837, .998,1686, .1748) /gene="Ert-2" /note="erythroid cell-specific and testis-specific gene transcript 2" /codon_start=1 /product="erythroid cell-specific and testis-specific protein 2 (ERT-2)" /protein_id="CA843518.1" /db_xref="GI:4902857" /translation="MPSRRGTSSNSRRSRNRNRRNSSSSCNSSNRNNSKKNKAKSN LQPGSGVGVCHSALICQGLHQGLHYL" 999, .1685 /gene="Ert-2" /number=1 1036, .1685 /gene="Ert-2" /number=1 1686, .2128 /gene="Ert-2" /number=2
CDS	
intron	
intron	
exon	
BASE COUNT	656 a 510 c 663 g 702 t 4 others
ORIGIN	
Query Match	84.0%: Score 16.8; DB 10; Length 2535;
Best local Similarity	90.0%: Pred. No. 1.6e+02;
Matches 18; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
Oy	1 gaacggccatcactgctgt 20
Db	1079 GAAGAGGCCATATCTGCTGT 1098
RESULT 12	
MMU242830	
LOCUS	MMU242830 4953 bp DNA ROD 04-OCT-1999
DEFINITION	Mus musculus Hemt gene, exons 1-3.
ACCESSION	AJ242830
VERSION	AJ242830.1 GI:6013070
KEYWORDS	GPI-anchored protein; Hemt gene; Hemt-3 protein.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. 1 (bases 1 to 4953) Xue,H., O'Neill,D., Morrow,D. and Bank A. A novel mouse gene, Hemt, encoding an hematopoietic cell-specific transcript Gene 231 (1-2), 49-58 (1999)
JOURNAL	
MEDLINE	99250249
REFERENCE	2 (bases 1 to 4953)
AUTHORS	Xue,H., O'Neill,D., Wang,X., Wolgemuth,D. and Bank A.
TITLE	Hemt3, encoding a 22 kDa polypeptide likely to be a GPI-anchored protein, is expressed in spermatocytes only unpublished 3 (bases 1 to 4953) Xue,H.

TITLE	Direct Submission			
JOURNAL	Submitted (02-JUN-1999) Xue H., Genetics and Development, Columbia University College of Physicians and Surgeons, Room 1614, HSC, 701, West 168th Street, New York, NY 10032, USA			
FEATURES	Location/Qualifiers			
source	1..4953			
	/organism="Mus musculus"			
	/db.xref="taxon:10090"			
mRNA	join(407..806,1385..1492,4103..4545)			
gene	/gene="HemT"			
	407..4545			
	/gene="HemT"			
exon	407..806			
	/gene="HemT"			
	/number=1			
CDS	join(680..806,1385..1492,4103..4398)			
	/gene="HemT"			
	/codon_start=1			
	/product="HemT-3 protein"			
	/protein_id="CA857315.1"			
	/db.xref="GI:6013071"			
	/translation="MMPEFLSLMGIPWVDNISINNSGLVFNQITGLDNIETPRM			
	POMCHSCOEHTFYCPRIHGYCDMDIRCLITAIKVNIRLIYKQCTKDTFIREM			
	VPEPLPVYKADKNRYFVMCCSITTCNVGPTNLREDLDETSEEVVARECLGWM			
	MLCEFLHLSLIIT"			
intron	807..1384			
	/gene="HemT"			
	/number=1			
exon	1385..1492			
	/gene="HemT"			
	/number=2			
intron	1493..4102			
	/gene="HemT"			
	/number=2			
exon	4103..4545			
	/gene="HemT"			
	/number=3			
BASE COUNT	1279 a	1018 c	1337 g	1310 t
ORIGIN	9 others			
Query Match	84.0%; Score 16.8; DB 10;			
Best Local Similarity	90.0%; Pred. No. 1.6e+02;			
Matches	18;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;
QY	1 gaacggccatacctgctgt 20			
Db	3497 GAAAGGCCATATCTGCTGT 3516			
RESULT 13				
AB000617				
LOCUS	AB000617	35000 bp	DNA	HCT
DEFINITION	Bacillus subtilis genomic DNA, 22 to 25 degree region, complete cds.			
ACCESSION	AB000617			
VERSION	AB000617.1	GI:2415716		
KEYWORDS	YccK; NaB; YceG; YceK; YceJ; YceI; YceL; YceH; YceN; YceG; YceF; YceE; YceD; YceD; YceB; YceA; YceI; YceH; YceG; YceF; RapJ; YcdJ; YcdC; YcdB; YcdA; YccK; YccH; YccG; YccF; Ycce; YccC; YccB; Trlacylglycerol lipase; LmrA; YccA; LmrB; YcdU.			
SOURCE	Bacillus subtilis (strain:168 tpc2) DNA.			
ORGANISM	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.			
REFERENCE	1 (bases 1 to 35000)			
AUTHORS	Yamane,K.			
TITLE	Direct Submission			
JOURNAL	Submitted (20-JAN-1997) to the DDBJ/EMBL/Genbank databases. Kunio Yamane, University of Tsukuba, Institute of Biological Sciences, 1-1-1 Ten-Odai, Tsukuba, Ibaraki 305, Japan (Tel:0298-53-6680, Fax:0298-53-6680)			

REFERENCE	2 (sites)
AUTHORS	Kumano,M., Tamakoshi,A. and Yamane,K.
TITLE	A 32 kb nucleotide sequence from the region of the lincomycin-resistance gene (22-25 degree) of the <i>Bacillus subtilis</i> chromosome and identification of the site of the lin-2 mutation unpublished (1997)
JOURNAL	
FEATURES	Location/Qualifiers
source	1. .35000
	/organism="Bacillus subtilis"
	/strain="168 trpC2"
	/db_xref="taxon:1423"
	/map="22 to 25 degree"
gene	592. .1704
	/gene="yccb"
CDS	592. .1704
	/gene="yccb"
	/note="homologue of class-V proteins of pyridoxal-phosphate-dependent aminotransferase"
	/codon_start=1
	/transl_table=1
	/product="yccb"
	/protein_id="BAA22227.1"
	/db_xref="GI:2415717"
	/translation="MEHLPEYROLPTLQHTMLASCSQALAEVSRALQDYVSL LYGTMMKEALERTERANNEFAKLGAEPDEVALVPSDALVSAVSLTARCKHY YTDMPFAPVHWQASHDYTVSIPIDGVLPLOYETHSDFTVLCVPHYHNDGY VODIKALAEISORSGSLFVDAVQSAGHPIIDVEMGVDMLAGTKYILGIPVAF VYREKLADALPKASAMPQSPGCAKAKARFPOTGTAFISVTAARALSLNHI GVSHRDHVKTCADAVQAAERKGLQAAAGGIGDPQWAVARD RASSETAGLLKKKY ICARENVIRLAPHFTYKEEMRAHIDIAKTIHK"
gene	complement(1746. .3185)
	/gene="ycca"
CDS	complement(1746. .3185)
	/gene="ycca"
	/note="lincomycin-resistance protein, similar to multidrug-resistant protein: lmrB gene"
	/codon_start=1
	/transl_table=1
	/product="lmrB"
	/protein_id="BAA22228.1"
	/db_xref="GI:2415718"
	/translation="MILETTAKASQYKWEIMISLLAGFIQMSSEALNIALTDLK KELNITATVQMLTGYLVGLIVPVSGLQWFTTRQFTVSLIFSLGFIATAA PSFLEHLARIVQALGTGLPLMFNTLIVFPBKGAMGTIGIVIMPAGIPNF SGVLEHLNMMHFEWISLPFLVAFVIAVQWVSSTPKRIDVSTIISTIGGSI VFGSNAGSGSGSPVYSLIVGVGILFISITQIMKOPMNLAEATYPTTIG VIVFICMWIISSMILPMYIOGGLVITRASSGLVLEGGINGMSVGRULPDI GPKULVPGVITVYVLMFSSVTTTSTAVLIIILHTCLMIGISMIMPADQNLNDY PREYPPGTAIMNTLOQMGALGTAVAVSIMAAQHDHMYSTVKNPADPAVIPAOLTAG VQRAVFPAMIVAILIGLIGAFPKRKVDH"
gene	complement(3225. .3791)
	/gene="yccb"
CDS	complement(3225. .3791)
	/gene="yccb"
	/note="lmrA gene; probable repressor of lincomycin-resistance operon"
	/codon_start=1
	/transl_table=1
	/product="lmrA"
	/protein_id="BAA22229.1"
	/db_xref="GI:2415719"
	/translation="MSYDSREKTIISAATRLPOLQGYGTGHNQIIKESGAPKGSLEY HFGCKEOLAEIVANDMEKEYIRKIDACMACCTDPAGIQAFLKELSCQSCFEDIG LPVGLAAETSLKSEPLREACHBAVKEWASVYBEKIKOTGCSSEKAKASVYVNAIE GILLISITAKNSPLIHISCIPLDLKR"
gene	4008. .5135
	/gene="yccb"
CDS	4008. .5135
	/gene="yccb"
	/note="probable L-asparaginase precursor"
	/codon_start=1
	/transl_table=1
	/product="yccb"
	/protein_id="BAA22230.1"
	/db_xref="GI:2415720"
	/translation="MKRQRLVFTLALFEYTCGSHSPETRESKPEKAGTOXVSASA SEKKDPNIRILATGCTIAGADSKISTPEYAGVAGVGSILAEVPEKMDIANVSEEO IIVAGSTNDNKKILAKLRINHLASDDVDGIVYHGTDLLETAFLILVKSXP VIVVSGMRPSTIASADGPSNLYNAVKVAGAPKAGKGLVLYLNDRIASARYVKTNTT TYLTFKSEEMGVGTIADDIYFNNEITRKTKDFTSVSNLDELPQVDIIYGYQNGS YLPDAVAKKAGAGIVFAGSGNGSLSDAEKAGASAKGVVIVSTRTGNGVYTPMD YAEKDLASNSLNPQKARMLMLALTKTNDPQIAQVNEY"
gene	5298. .5936
	/gene="lipA"
CDS	5298. .5936
	/gene="lipA"
	/note="lipA"
	/codon_start=1
	/transl_table=1
	/product="Triacylglycerol lipase"
	/protein_id="BAA22231.1"
	/db_xref="GI:2415721"
	/translation="MKFVRRIIALVITLMLSVTSLSLALOPSARAAHNPVAVHGIG GASFNPAGIKSTIVSOGSRBDKLYAVDFMDKGTGYNNGPVLSPFDQDETGAKV DIVAHSMGAGNTLYIKNLGDKNKVANVYVLTGAGANRLPTGALPDTQPKDILYTSIY SSADMIYMYNYSRLDGAARNQIHGVGHIGLVSQVNSLIEGINSQGQNTN"
gene	6073. .6318
	/gene="yccf"
CDS	6073. .6318
	/gene="yccf"
	/note="yccf"
	/codon_start=1
	/transl_table=1
	/product="yccf"
	/protein_id="BAA22232.1"
	/db_xref="GI:2415722"
	/translation="MHPRYQVTDNGLRGFPFICFPHDLOPDQOFFAOCALPICORHQ QRYHMQYQDNDNDTQSPHAKRSQCQNHCNDYIK"
gene	6592. .7668
	/gene="yccf"
CDS	6592. .7668
	/gene="yccf"
	/note="yccf"
	/codon_start=1
	/transl_table=1
	/product="yccf"
	/protein_id="BAA22233.1"
	/db_xref="GI:2415723"
	/translation="MGVKKRNACPCGSGSKTKKCCGSKVDPFPAELAKEAKQIOED LVEVAFYHRSISGFIHQHDFLSAMDROFKDISENGLMGIFEPHLAGERTIFBEY LOKKDSDITRPTREIYESWQSTPALILKDLKGLIHPEDVITAKOFEVEDMSNQ DLPYVGSILGYPITHEAKAEFPMTTIPPVRTALISKVKYKAAVAKDGTPEDF MKQEFNVILFALLAEKDEPOAKAEVTEMANDLEKETAALIEGMSGEEYPTIELI PAVIDIKTFPEKSPYIRKPEKFAAAYVYNALSLNGASVQAKLAKKYVSASTI SSRYEITSTQDPAERFQALSS"
gene	complement(7708. .8664)
	/gene="yccg"
CDS	complement(7708. .8664)
	/gene="yccg"
	/note="yccg"
	/codon_start=1
	/transl_table=1
	/product="yccg"
	/protein_id="BAA22234.1"
	/db_xref="GI:2415724"
	/translation="MILFQCLYLLFESFICYGQAAAFSSHTASVLAALGAAGAAGL YTMNTRVYKHSKSGGLCAMIAYOVYSGVYLIGDIMPVLYAIIFACCEGLRAGS ALQARTLSQIDKLTLOAHQAHNMGLDPSRNHDPKRTITAIKSAQPKADTOYIONMA DOYSOYRFLKGENAYAGVLDPLEKASVNSVSIHMTPLSPSPRAQVSLVG NILEALDASAEAREKAEIRLFTSLRSGLYVLTENSTPQMDKRVLDITVOSFGRSTK NGAHEGMGTYIIQGLYKGAEGRLDFTYBHPIFRLEIKIPROK"
gene	complement(18677. .9378)
	/gene="yccH"
CDS	complement(18677. .9378)
	/gene="yccH"
	/note="homologue of sensor/response regulator protein of <i>Staphylococcus aureus</i> "
	/codon_start=1
	/transl_table=1
	/product="yccH"


```

repeat_region 14324..14357
/note="TIGER1 repeat: matches 24..57 of consensus"
repeat_region 14358..14679
/note="AluX repeat: matches 1..309 of consensus"
repeat_region 14680..15805
/note="TIGER1 repeat: matches 57..1183 of consensus"
repeat_region 17081..17182
/note="51 copies 2 mer tt 60% conserved"
repeat_region 17126..17200
/note="15 copies 5 mer ttata 70% conserved"
repeat_region 17888..17921
/note="17 copies 2 mer gt 97% conserved"
repeat_region 18534..18646
/note="MER58 repeat: matches 2297..2419 of consensus"
repeat_region 18673..18768
/note="MER58 repeat: matches 75..179 of consensus"
repeat_region 18906..18986
/note="MIR repeat: matches 50..131 of consensus"
repeat_region 19039..19115
/note="L2 repeat: matches 2591..2672 of consensus"
repeat_region 21322..21622
/note="AluSX repeat: matches 1..305 of consensus"
repeat_region 22912..23027
/note="L2 repeat: matches 2581..2705 of consensus"
repeat_region 23877..24057
/note="FRAM repeat: matches -2..167 of consensus"
repeat_region 25919..26111
/note="MIR repeat: matches 18..215 of consensus"
repeat_region 26649..27092
/note="LIM4 repeat: matches 704..1090 of consensus"
repeat_region 27336..27468
/note="MER5A repeat: matches 18..144 of consensus"
repeat_region 27838..28233
/note="MER4B repeat: matches 192..564 of consensus"
repeat_region 28244..28543
/note="AluSX repeat: matches 1..292 of consensus"
repeat_region 28556..28678
/note="MER4D repeat: matches 88..208 of consensus"
repeat_region 29026..29087
/note="MIR repeat: matches 84..141 of consensus"
repeat_region 29351..29600
/note="MIR repeat: matches 10..259 of consensus"
repeat_region 29601..29820
/note="match: GSS: Em:A0601481"
misc_feature 29741..29774
/note="17 copies 2 mer ac 100% conserved"
repeat_region 30325..30496
/note="MSTP repeat: matches 218..392 of consensus"
repeat_region 30556..30787
/note="AluSG repeat: matches 1..233 of consensus"
repeat_region 30833..30932
/note="MTR1A2 repeat: matches 1..102 of consensus"
repeat_region 31028..31319
/note="AluJ repeat: matches 7..297 of consensus"
repeat_region 31876..31915
/note="20 copies 2 mer gt 100% conserved"
misc_feature 31916..32945
join(<32945..33158,62655..62764,68856..69656,76116..76409,
76945..77048,77885..78070,81805..81988,91988..92113,
92334..92539,93956..94129,95108..95137,96248..96468,
97986..97852,105801..105892,107740..107908,110501..110629,
111177..111273,116309..118340)
/gene="dJ683H8.1"
/note="match: CDNAS: Em:AF081150 Em:AF081151 Em:AF081152
Em:AF081153 Em:AF111069 Em:AF081154 Em:AF081155
Em:AF081156 Em:AF111083 Em:AF111082 Em:AF081157
Em:AF081158 Em:AF081159 Em:AF081145 Em:AF081149
Em:AF081148 Em:U78105 Em:AF111081 Em:AF111070 Em:AF081147
Em:AF111071 Em:AF063102 Em:AF111072 Em:AF111080
Em:AF063103 Em:AF111073 Em:AF104939 Em:AF111074
Em:AF111075 Em:AF111076 Em:AF111077 Em:AF111078
Em:AF081144 Em:AF111079 Em:AF111099 Em:AF081146 Em:U72487
Em:AJ131581 Em:AF111084 Em:AF111088 Em:AF111090
Em:AF111085 Em:AF111092 Em:AF111086 Em:AF111087

```

MRNA

```

Em:AF111091 Em:AF111088 Em:AF111089 Em:AB020628
Em:AF111097 Em:AF111096 Em:AF111093 Em:AB018311
Match: ESTs: Em:AW320731 Em:T47903 Em:AT741155
Em:AW607395 Em:AL040055 Em:AA843854 Em:RI9057 Em:AW493841
Em:AA010286 Em:AL040055 Em:AA334422 Em:AA652670 Em:RS9662
Em:AA104027 Em:AT450192 Em:AM163402 Em:AT127338
Em:AA020067 Em:AW300567 Em:AA292171 Em:AW536248 Em:243013
Em:AT390472 Em:240660 Em:AA227775 Em:AA230841 Em:AW553551
Em:AA227776 Em:AW472715 Em:AT640464 Em:AW553551
Em:AW157344 Em:AA096135 Em:AA491424 Em:AM008295
Em:AA274510 Em:AT525897 Em:AA659353 Em:R27783 Em:AW023001
Em:AA058258 Em:T39186 Em:AT806350 Em:AT555886 Em:AA267884
Em:DB1314 Em:AT743567 Em:W10208 Em:AT390286 Em:R26636
Em:AW38191 Em:AA103165 Em:AT286770 Em:AT238196
Em:AA240842 Em:AA032925 Em:R27872 Em:AT970050 Em:AT805316
Em:AT568053 Em:AT991742 Em:H59160 Em:AW070628 Em:T92217
Em:NA8811 Em:AT180282 Em:T78230 Em:N88253 Em:AA166267
Em:AA36634 Em:AA48931 Em:R70738 Em:AT889659 Em:AT431621
Em:AA5605 Em:AT196359 Em:AA595513 Em:AT394405 Em:AT633760
Em:244892 Em:AT380727 Em:AT765930 Em:W9544 Em:W4533
Em:AA829215
/evidence=not-experimental
/product="dJ683H8.1.3 (lectomedin 1 gamma (LEC1, LPHR1))"
join(<32945..33158,62655..62764,68856..69656,76116..76409,
76945..77048,77885..78070,81805..81988,91988..92113,
92334..92539,93956..94129,95108..95137,96248..96468,
97786..97852,105801..105892,107740..107908,110501..110629,
116309..118340)
/gene="dJ683H8.1"
/note="match: CDNAS: Em:AF081150 Em:AF081151 Em:AF081152
Em:AF081153 Em:AF111069 Em:AF081154 Em:AF081155
Em:AF081156 Em:AF081148 Em:AF081147 Em:AF081157
Em:AF081158 Em:AF081159 Em:AF081146 Em:U78105 Em:AF081145
Em:AF081144 Em:AF111070 Em:AF081149 Em:AF111080
Em:AF04938 Em:AF111071 Em:AF063102 Em:AT11072
Em:AJ131581 Em:AF063103 Em:AF111073 Em:U72487 Em:AF111074
Em:AF111075 Em:AF111076 Em:AF111077 Em:AF111078
Em:AF111079 Em:AF111099 Em:AF111084 Em:AF111083
Em:AF111082 Em:AF111081 Em:AF111090 Em:AF111091
Em:AF111096 Em:AF111092 Em:AF111097 Em:AF111093
Em:AB020628 Em:AF111085 Em:AF111086 Em:AF111087
Em:AF111098 Em:AF111088 Em:AF111089 Em:AB018311
Match: ESTs: Em:AW320731 Em:T47902 Em:AW493841 Em:T47903
Em:AT741155 Em:AA843854 Em:AT400055 Em:RI9057 Em:AA705981
Em:AW607395 Em:AA334422 Em:AA553551 Em:AA096135
Em:AA652670 Em:AT431621 Em:RS9662 Em:AW491424 Em:AT450192
Em:AW008295 Em:AT127338 Em:AT525897 Em:AA300567
Em:AW292171 Em:AW536248 Em:243013 Em:T92217 Em:240660

```

```

Query Match 84.0%; Score 16.8; DB 9; Length 146300;
Best Local Similarity 90.0%; Pred. No.2e+02;
Matches 18; Conservative 2; Indels 0; Gaps 0;
QY 1 gaaagcgcaacactgctgt 20
|||||
Db 107523 GAAAGCGCACTGCTGT 107542

```

Search completed: April 3, 2002, 11:00:10
Job time: 5201 sec

Thu Apr 4 09:39:28 2002

us-09-885-799-488.rge

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 11:05:29 ; Search time 321.05 Seconds
(without alignments)
53,408 Million cell updates/sec

Title: US-09-885-799-488

Sequence: 1 gaacagcctaccctgctgt 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

N_Geneseq_1101:*

1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length DB	ID	Description
1	20	100.0	3283	20	AAZ31789
2	16.8	84.0	394	17	AAZ44707
3	16.8	84.0	394	18	AAZ78143
4	15.8	79.0	381	21	AAZ11711
5	15.8	79.0	486	21	AAZ43490
6	15.4	77.0	1533	21	AAZ42159
7	15.4	77.0	2990	21	AAZ79597
8	15.4	77.0	109973	21	AAZ22298
9	15.2	76.0	256	14	AAZ60899
10	15.2	76.0	945	22	AAZ32333
11	15.2	76.0	1654	22	AAZ58720

12	15.2	76.0	2015	21	AAZ98156
13	15.2	76.0	2055	22	AAZ07750
14	15.2	76.0	2403	22	AAZ62743
15	15.2	76.0	2407	15	AAZ070421
16	15.2	76.0	2456	21	AAZ49194
17	15.2	76.0	2534	22	AAZ08353
18	15.2	76.0	2573	22	AAZ158719
19	15.2	76.0	2879	22	AAZ07718
20	15.2	76.0	3420	22	AAZ160505
21	15.2	76.0	3420	22	AAZ160505
22	15.2	76.0	6312	18	AAZ90691
23	15.2	76.0	8088	21	AAZ45540
24	15.2	76.0	8535	15	AAZ073731
25	15.2	76.0	8625	18	AAZ45143
26	15.2	76.0	8625	19	AAZ40146
27	15.2	76.0	11233	19	AAZ40151
28	15.2	76.0	43804	18	AAZ68375
29	15.2	76.0	43804	20	AAZ26690
30	15.2	76.0	44018	22	AAZ82392
31	15.2	76.0	56583	21	AAZ21125
32	15.2	76.0	56583	21	AAZ35003
33	15.2	76.0	305107	22	AAZ62689
34	15.2	75.0	7035	22	AAZ163778
35	15	75.0	7035	22	AAZ163778
36	14.8	74.0	120	21	AAZ08558
37	14.8	74.0	858	22	AAZ05464
38	14.8	74.0	1798	15	AAZ06590
39	14.8	74.0	2037	15	AAZ078149
40	14.8	74.0	2113	22	AAZ15875
41	14.8	74.0	2428	21	AAZ87190
42	14.4	72.0	355	19	AAZ20246
43	14.4	72.0	439	22	AAZ82494
44	14.4	72.0	503	22	AAZ82772
45	14.4	72.0	602	21	AAZ03857

ALIGNMENTS

RESULT 1	AAZ31789
ID	AAZ31789 standard; DNA; 3283 BP.
XX	
AC	AAZ31789;
XX	
DT	21-JAN-2000 (first entry)
XX	
DE	HPV70 DNA sequence.
XX	
KW	HPV68: human papillomavirus 68; detection; immunisation; infection;
XX	HPV70; ss.
OS	Human papillomavirus.
XX	
PN	US5981173-A.
XX	
PD	09-NOV-1999.
XX	
PF	11-FEB-1997; 97US-0815667.
XX	
PR	14-FEB-1996; 96US-0011650.
XX	
PR	15-FEB-1996; 96US-0020458.
XX	
PA	(INSP) INST PASTEUR.
XX	
PI	(INRM) INST NAT SANTE & RECH MEDICALE.
XX	
DR	Longuet M., Orth G., Beaudenon S.
XX	
WP	WPI: 1999-633305/54.
XX	
PT	New human papillomavirus (HPV) 68 and 70 DNA sequences, useful in the
XX	treatment of HPV infections -

PS Disclosure: Fig 6; 25pp; English.

XX This sequence represents a human papillomavirus (HPV) 70 DNA sequence.

CC The invention relates to a purified HPV68 DNA (I) deposited at C.N.C.M.

CC (Collection Nationale de Culture de Microorganismes) under Accession No.

CC 1-1540, and a HPV70 DNA (II) sequence. Fragments of the HPV68 and HPV70

CC DNA may be useful as primers or probes in the detection of HPV68 or HPV70

CC infections. Proteins produced by the HPV DNA's can be used to immunise

CC animals against HPV68 or HPV70 infections.

SQ Sequence 3283 BP; 945 A; 640 C; 699 G; 999 T; 0 other;

QY 1 gaacgagccatcctcgt 20

DB 1070 gaacgagccatcctcgt 1089

Query Match 100.0%; Score 20; DB 20; Length 3283;

Best Local Similarity 100.0%; Pred. No. 0.42;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2

AA14707

ID AA14707 standard; DNA: 394 BP.

XX

AC AA14707;

XX

DT 29-JAN-1997 (first entry)

XX

DE Human papillomavirus isolate Pap16 L1 region sequence.

XX

KW Probe: primer: PCR: polymerase chain reaction: amplification:

KW human papillomavirus; consensus: ss.

XX

OS Human papillomavirus.

XX

PN US5527898-A.

XX

PD 18-JUN-1996.

XX

PF 09-SEP-1988; 88US-0243486.

XX

PR 24-SEP-1993; 93US-0126452.

PR 09-SEP-1988; 88US-0243486.

PR 10-MAR-1989; 89US-0322550.

PR 09-SEP-1988; 89MO-US03747.

PR 14-NOV-1990; 90US-0613142.

PR 20-APR-1993; 93US-0050743.

PR 07-JUN-1995; 95US-0474542.

XX

PA (HOFF) HOFFMANN LA ROCHE INC.

XX

PI Bauer HM, Gravitt PE, Greer CE, Manos MM, Resnick RM;

PI Zhang TY;

XX

DR WPI: 1996-299903/30.

XX

PT Nucleic acid hybridisation probes - specific for selected human

PT papilloma virus types

XX

PS Claim 19: Column 23-24; 96pp; English.

XX

CC The invention relates to new oligonucleotide probes and primers used

CC for the detection of human papillomaviruses (HPV) which are not genital

CC types 6, 11, 16, 18 or 33. The probes and primers AA144608-T44693 are

CC esp. used to detect HPV types 26, 31, 31B, 35, 39, 40, 43, 45, 51-59 and

CC 68. The primers can be used to detect these HPV types in conjunction with

CC the consensus primers and typing probes AA144733-T44906, which are based

CC on and amplify fragments of the L1, E6, E7 and E1 regions of the HPV

CC sequences. The use of consensus probes and primers also allowed the

CC isolation of novel HPV isolates and strains. This sequence represents

CC the L1 region of isolate Pap16.

XX

SQ Sequence 394 BP; 127 A; 80 C; 74 G; 113 T; 0 other;

QY 1 gaacgagccatcctcgt 20

DB 41 gaacgagccatcctcgt 60

Query Match 84.0%; Score 16.8; DB 17; Length 394;

Best Local Similarity 90.0%; Pred. No. 16;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 3

AA178143

ID AA178143 standard; DNA: 394 BP.

XX

AC AA178143;

XX

DT 07-OCT-1997 (first entry)

XX

DE DNA sequence of L1 amplified region of human papillomavirus Pap16.

XX

KW L1: amplified region: HPV: probe: detection; typing: ss.

XX

OS Human papillomavirus Pap16.

XX

PN US5639871-A.

XX

PD 17-JUN-1997.

XX

PF 09-SEP-1988; 88US-0243486.

XX

PR 24-SEP-1993; 93US-0126452.

PR 09-SEP-1988; 88US-0243486.

PR 10-MAR-1989; 89US-0322550.

PR 09-SEP-1988; 89MO-US03747.

PR 14-NOV-1990; 90US-0613142.

PR 20-APR-1993; 93US-0050743.

PR 01-JUN-1995; 95US-0457648.

XX

PA (HOFF) ROCHE MOLECULAR SYSTEMS INC.

XX

PI Bauer HM, Gravitt PE, Greer CE, Imprial CC, Manos MM;

PI Resnick RM, Zhang TY;

XX

DR WPI: 1997-332084/30.

XX

PT New oligonucleotide probes for human papilloma-virus - used for

PT detecting and typing HPV and for detecting previously unknown HPV

PT types and subtypes

XX

PS Disclosure: Columns 169-170; 94pp; English.

XX

CC The present sequence is the L1 amplified region of human

CC papillomavirus Pap16 (HPVpap16), useful as a probe to detect and

CC type HPV in samples. It can also be used to detect previously

CC unknown HPV types and subtypes.

XX

SQ Sequence 394 BP; 127 A; 80 C; 74 G; 113 T; 0 other;

QY 1 gaacgagccatcctcgt 20

DB 41 gaacgagccatcctcgt 60

Query Match 84.0%; Score 16.8; DB 18; Length 394;

Best Local Similarity 90.0%; Pred. No. 16;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 4

AA11711/C

ID AAF11711 standard; cDNA; 381 BP.
XX
AC AAF11711;
XX
DI 13-MAR-2001 (first entry)
XX
DE Aspergillus niger EST SEQ ID NO:4234.
XX
KW Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus niger.
XX
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US07781.
XX
PR 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PA (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB,
DR WPL: 2000-594572/56.
XX
XX Monitoring differential expression of genes in filamentous fungal cells
PT using fluorescence-labeled nucleic acids isolated from the cells and a
XX substrate of expressed sequence tags -
XX
PS Claim 87: Page 1840; 31c1p; English.
XX
CC The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as or spot on an
CC array equals one gene or open reading frame, and Organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from
CC Aspergillus niger; AAF11854 to AAF14878 represents ESTs from Aspergillus
CC niger; AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC all specifically claimed in the present invention.
XX
XX Sequence 381 BP; 88 A; 87 C; 95 G; 111 T; 0 other;

Query Match 79.0%; Score 15.8; DB 21; Length 381;
Best Local Similarity 89.5%; Pled. No. 53;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaacggccatactgctg 19
Db 79 GACACGGCCATCCCTGCTG 61

RESULT 5
AAC43490
ID AAC43490 standard; DNA; 486 BP.
XX
AC AAC43490;
XX
DI 17-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 39429.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
XX Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137228.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143277.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144486.
PR 19-JUL-1999; 99US-0144487.
PR 19-JUL-1999; 99US-0144431.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.0%; Score 15.8; DB 21; length 486;
Best Local Similarity 89.5%; Pred. No. 54;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaaacggcataactcgtg 19
I | | | | | | | | | | |
Db 190 gaaacggcataactcgtg 208

RESULT 6
AAC42159/c
ID AAC42159 standard; DNA; 1533 BP.
XX
AC AAC42159;
XX

DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 34506.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
PD
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130444.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131444.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132487.
PR 05-MAY-1999; 990S-0132488.
PR 06-MAY-1999; 990S-0132488.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 14-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137558.
PR 04-JUN-1999; 990S-0137552.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 14-JUN-1999; 990S-0138847.
PR 16-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 17-JUN-1999; 990S-0139453.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140355.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 02-JUL-1999; 990S-0142194.
PR 06-JUL-1999; 990S-0142055.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 20-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145613.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.

```

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161407.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 77.0%; Score 15.4; DB 21; Length 1533;
Best Local Similarity 94.1%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 gaaacgacatactgc 17
    |||||
Db 1350 GAACGCGCATAGCTGC 1334

```

```

RESULT 7
AAC79597/c
ID AAC79597 standard; DNA: 2990 BP.
XX
AC AAC79597;
XX
DT 08-FEB-2001 (first entry)
XX
DE Virulence gene #17.
XX
KM Virulence gene; antibacterial; vaccine; bacterial infection;
KW septicemia; bronchopneumonia; rhinitis; wound infection; ss.
XX

```

```

OS Pasteurella multocida.
XX
XX WO200061724-A2.
XX
XX 19-OCT-2000.
XX
XX 06-APR-2000; 2000WO-US09218.
XX
XX 09-APR-1999; 99US-0128669.
XX 10-SEP-1999; 99US-0153453.
XX
XX (PHAA ) PHARMACIA & UPJOHN INC.
XX
XX Lowery DE, Fuller TE, Kennedy MT;
XX
XX WPI: 2000-647422/62.
XX
XX P-PSDB; AAB44537.
XX
XX Attenuated Pasteurellaceae bacteria comprising mutations in virulence
XX genes, useful as a live attenuated vaccine against bacterial infections
XX
XX Claim 7; Pages 144-146; 322pp; English.
XX
XX The family Pasteurellaceae encompasses several pathogens that infect a
XX wide variety of animals. The present invention relates to virulence genes
XX from Pasteurellaceae. The present sequence is one such virulence gene.
XX The present sequence may be mutated in order to produce an inactive gene.
XX The inactive virulence gene may in turn be used to produce a vaccine,
XX which is useful for treating bacterial infections such as septicemias,
XX bronchopneumonias, rhinitis and wound infections.
XX
XX Sequence 2990 BP; 1012 A; 465 C; 648 G; 865 T; 0 other:

```

```

Query Match 77.0%; Score 15.4; DB 21; Length 2990;
Best Local Similarity 94.1%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 gaaacgacatactgc 17
    |||||
Db 2393 GAACGCGCATAGCTGC 2377

```

```

RESULT 8
AAF22298/c
ID AAF22298 standard; DNA: 109973 BP.
XX
AC AAF22298;
XX
XX 20-MAR-2001 (first entry)
XX
DE BAC containing repeats from centromeres 1-4 #21.
XX
XX Centromere; michrosome; vector; ds.
XX
XX Arabidopsis thaliana.
XX
XX WO20005325-A2.
XX
XX 21-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-US07392.
XX
XX 18-MAR-1999; 99US-0125219.
XX 01-APR-1999; 99US-0127409.
XX 18-MAR-1999; 99US-0134770.
XX 13-SEP-1999; 99US-0153584.
XX 17-SEP-1999; 99US-0154603.
XX
XX (UVCH-) UNIV CHICAGO.
XX
XX Preuss D, Copenhaver G, Keith K;

```

XX
DR WPI: 2000-587529/55.
XX
PT Recombinant DNA construct comprising a plant centromere, useful for
PT producing stably inherited chromosomes which can serve as vectors for
PT the construction of transgenic plant and animal cells -
XX
PS Claim 102; Page 738-763; 1449pp; English.
XX
CC The present invention relates to a recombinant DNA construct of a plant
CC (Arabidopsis thaliana) centromere. The constructs are useful for
CC producing stably inherited chromosomes which can serve as vectors for
CC the construction of transgenic plant and animal cells expressing
CC selected proteins such as hormones, enzymes, interleukins, clotting
CC factors, cytokines, antibodies, and growth factors.
XX
SQ Sequence 109973 BP; 31657 A; 23911 C; 22655 G; 31745 T; 5 other:

Query Match 77.0%; Score 15.4; DB 21; Length 109973;
Best Local Similarity 94.1%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gaacggccatctgtc 17
|||||
DB 58921 GAACGGCCATCTGTC 58905

RESULT 9
AA060899/C
ID AA060899 standard; DNA; 256 BP.
XX
XX AA060899:
AC 16-MAR-1994 (first entry)
XX
DT 16-MAR-1994 (first entry)
XX
DE Human brain Expressed Sequence Tag EST00990.
XX
KW Gene transcription product; genetic markers; tagging; in vivo;
KW transcription; mapping; locations; chromosomes; chromosomal; ss.
XX
OS Homo sapiens.
XX
PN W09316178-A.
XX
PD 19-AUG-1993.
XX
PF 12-FEB-1993; 93WO-US01294.
XX
PS 12-FEB-1992; 92US-0837195.
XX
PR 12-FEB-1992; 92US-0837195.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Adams MD, Moreno RF, Venter CJ;
XX
DR WPI: 1993-272882/34.
XX
PT Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
PT of most human genes
XX
PS Example 4; Page 408; 500pp; English.
XX
CC The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for human genes transcribed in vivo. They can be used to facilitate
CC tagging of most human genes, for mapping locations of expressed genes
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prepn. of antisense sequences, probes and constructs.
CC EST00990 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM. See also AA059041-061440.
XX

SQ Sequence 256 BP; 89 A; 53 C; 43 G; 71 T; 0 other:

Query Match 76.0%; Score 15.2; DB 14; Length 256;
Best Local Similarity 85.0%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gaacggccatctgtc 20
|||||
DB 235 GAACGGCCATCTGTC 216

RESULT 10
AAH32333/C
ID AAH32333 standard; DNA; 945 BP.
XX
XX AAH32333:
AC 30-JUL-2001 (first entry)
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 906.
XX
KW Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN W0200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000MO-US27582.
XX
PS 08-OCT-1999; 99US-0158615.
XX
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI: 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 541-542; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 945 BP; 186 A; 257 C; 208 G; 294 T; 0 other:

Query Match 76.0%; Score 15.2; DB 22; Length 945;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gaacggccatctgtc 20
|||||

Db 612 GAACAGGCATACATGCTGT 593

RESULT 11

AA158720/c

ID AA158720 standard; cDNA; 1654 BP.

XX AA158720;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 923.

DE

XX Human; nootropic; immunosuppressant; cytoskeletal; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokine;

KW chemokine; thrombolytic; drug screening; arthritis; inflammation;

KW leukemia; ss.

XX Homo sapiens.

XX W0200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620112.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RJ;

XX WPI: 2001-442253/47.

DR P-PSDB; AAM39564.

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries .

PS Claim 1: SEQ ID NO 923; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and

CC the encoded polypeptides (AAM8642-AAM42213) with nootropic,

CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activity, chemokine/chemokine activity, haemostatic

CC activity/inhibin activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX Sequence 1654 BP; 406 A; 404 C; 485 G; 359 T; 0 other;

XX Query Match 76.08; Score 15.2; DB 22; Length 1654;

XX Best Local Similarity 85.08; Pred. NO. 1.3e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gaacggcatatgctgt 20

|||||

Db 817 GAACGGCATATGCGCACT 798

RESULT 12

AA298156

ID AA298156 standard; cDNA; 2015 BP.

XX AA298156;

XX 11-MAY-2000 (first entry)

DE Human signal peptide containing protein HSP-48 cDNA SEQ ID NO:182.

KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;

KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;

KW antimicrobial; nootropic; neuroprotective; cardiovascular; hepatotropic;

KW antiasthmatic; gene therapy; cell proliferation; neurological disorder;

KW reproductive disorder; developmental disorder; arteriosclerosis;

KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;

KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;

KW Parkinson's disease; Huntington's diseases; ovulatory defect;

KW muscular dystrophy; ss.

XX Homo sapiens.

XX W0200000610-A2.

XX 06-JAN-2000.

XX 25-JUN-1999; 99MO-US14484.

XX 26-JUN-1998; 98US-0090762.

XX 31-JUL-1998; 98US-0094983.

XX 01-OCT-1998; 98US-0102686.

XX 11-DEC-1998; 98US-0112129.

XX (INCY-) INCYTE PHARM INC.

PI Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;

PI Aketplom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;

PI Bandman OJ;

XX WPI: 2000-160673/14.

DR P-PSDB; AAY87271.

XX New human signal peptide-containing proteins useful in treatment,

PT prevention and diagnosis of e.g. cancer, inflammation and

PT cardiovascular disease .

PS Claim 9: Page 279-280; 327pp; English.

XX AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the

CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have

CC anticancer, anti-inflammatory, antimicrobial, nootropic, hepatotropic,

CC neuroprotective, cardiovascular and antiasthmatic activities, and can

CC be used in gene therapy. HSPs can be used to treat or prevent disorders

CC associated with decreased activity or function of HSP. Antagonists of

CC HSP are used to treat or prevent disorders associated with increased

CC activity or function of HSP. Such diseases include cell proliferation

CC (including cancer), inflammation, cardiovascular, neurological,

CC reproductive or developmental disorders, (e.g. arteriosclerosis,

CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,

CC asthma, Crohn's disease, microbial or other infections, congestive or

CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's

CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP

CC nucleic acids can be used for the recombinant production of HSP, for

CC detecting HSP in standard hybridisation and amplification assays (for

CC diagnosis and monitoring), in gene therapy, as antisense,

CC triplex-forming or ribozyme therapeutics, for detecting related sequences

XX
DR WPI: 2001-329085/34.
DR P-PSDB: AAE03337.

FF 06-NOV-2000; 2000MOC-0528888.
 XX
 PR 24-NOV-1999; 99CN-0124717.
 XX

PA (PENT-) PE CORP NY.
 PA (THIR-) THIRD INST OCEANOGRAPHY STATE OCEANI C A.
 PA (SINO-) SINOGENOMAX CO LTD.
 XX
 PI Xu X, Yang F, He J, Pham L, He M, Ye Y, Shen Y, Kodira C;
 DR MPI; 2001-355877/37.
 DR P-PSDB: AAG84963.
 XX
 PT Primary nucleotide sequence of the shrimp white spot Bacilliform virus
 PT (WSBV), useful for producing viral polypeptides that can be used to
 PT screen for agents that are useful for treating WSBV infection -
 XX
 PS Claim 4: Figure 3: 626bp; English.
 XX
 CC The invention provides the primary nucleotide sequence of the WSBV genome
 CC (AAH62689), predicted transcript sequences (AAH62689-AAH62839) and
 CC encoded proteins (AAG84910-AAG85051) and oligonucleotide sequences
 CC (AAH62840-63160) suitable for use as primers or probes. The nucleic acid
 CC molecules and proteins of the invention are useful for diagnosis and
 CC monitoring viral infection. In screens for antiviral agents and for
 CC monitoring viral gene expression or activity during a treatment regimen.
 CC The nucleic acid molecules are also useful as antisense constructs to
 CC control viral gene expression in infected cells and tissues and to create
 CC transgenic viral resistant shrimp.
 CC
 SO Sequence 2403 BP; 797 A; 500 C; 481 G; 625 T; 0 other;

Query Match 76.0%; Score 15.2; DB 22; Length 2403;
 Best Local Similarity 85.0%; Pred. No. 1.4e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gaacggccatctcgt 20
 Db 1693 gaatggccatctcctat 1712

RESULT 15
 AA070421
 ID AA070421 standard: cDNA to mRNA; 2407 BP.
 XX
 AC AA070421:
 XX
 DT 27-FEB-1995 (first entry)
 XX
 DE C1C-K1 gene.
 XX
 KM C1C-K1; kidney; Henle's loop; probe; detection; chloride;
 KM diagnosis; ss.
 OS
 OS Rattus rattus - Sprague-Dawley.
 XX
 FH Key Location/Qualifiers
 FT CDS 44..2104
 FT /*tag= a
 XX
 XX JP06165680-A.
 PN
 XX
 PD 14-JUN-1994.
 XX
 PF 01-DEC-1992: 92JP-0343609.
 XX
 PR 01-DEC-1992: 92JP-0343609.
 XX
 PA (MARU/) MARUMO F.
 DR
 DR WPI; 1994-230226/28.
 DR P-PSDB: AAR60336.
 XX
 PT C1C-K1 cDNA probe specific for kidney Henle's loop kinked
 PT ascending limb maldistributed chloride channel RNA - useful in
 PT diagnosis of diseases caused by abnormal chloride transition

XX
 PS Claim 7: Page 7-10; 14pp; Japanese.
 XX
 CC The C1C-K1 probes (AA070422-23) are used in the detection of
 CC C1C-K1 DNA. Detection is useful in diagnosis of diseases
 CC caused by abnormal transition of chloride.
 XX
 SO Sequence 2407 BP; 472 A; 739 C; 655 G; 541 T; 0 other;

Query Match 76.0%; Score 15.2; DB 15; Length 2407;
 Best Local Similarity 85.0%; Pred. No. 1.4e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 gaacggccatctcgt 20
 Db 1621 gaacggccatctcgt 1640

Search completed: April 3, 2002, 11:05:32
 Job time: 5523 sec

Thu Apr 4 09:39:28 2002

us-09-885-799-488.rng

Page 11

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 10:31:11 ; Search time 143.87 Seconds
(without alignments)
31,484 Million cell updates/sec

Title: US-09-885-799-488

Sequence: 1 gaaagcgccatactgtctgt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexl 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	1059	2	US-08-815-667-11
2	20	100.0	1521	2	US-08-815-667-10
3	20	100.0	3283	2	US-08-815-667-14
4	16.8	84.0	394	1	US-08-474-542A-279
5	16.8	84.0	394	1	US-08-457-648-279
6	15.2	76.0	6312	4	US-09-011-745-7
7	15.2	76.0	8535	3	US-08-716-351A-1
8	15.2	76.0	8625	4	US-08-980-832-1
9	15.2	76.0	11233	4	US-08-980-832-27
10	14.8	74.0	2037	2	US-08-535-276-1
11	14.4	72.0	4651	2	US-08-449-644-4
12	14.4	72.0	4651	2	US-08-087-244A-4
13	14.2	71.0	618	4	US-08-961-810-17
14	14.2	71.0	618	4	US-08-352-902D-17
15	14.2	71.0	1290	4	US-09-293-395-6
16	14.2	71.0	1290	4	US-09-668-648-6
17	14.2	71.0	1320	4	US-09-293-395-14
18	14.2	71.0	1320	4	US-09-668-648-14
19	14.2	71.0	1461	3	US-08-445-638-5
20	14.2	71.0	1461	3	US-08-445-664C-5
21	14.2	71.0	1461	5	PCT-US94-03437-5
22	14.2	71.0	1671	2	US-08-662-560-1
23	14.2	71.0	1671	2	US-08-780-749A-5
24	14.2	71.0	1671	3	US-08-706-281A-15
25	14.2	71.0	1671	4	US-09-097-231-15
26	14.2	71.0	1671	4	US-08-870-511-5
27	14.2	71.0	1801	5	PCT-US95-02455-1

C 28	14.2	71.0	2164	4	US-09-318-794A-1	Sequence 1, Appl
C 29	14.2	71.0	2164	4	US-09-318-793A-3	Sequence 3, Appl
C 30	14.2	71.0	2369	2	US-08-883-534-4	Sequence 4, Appl
C 31	14.2	71.0	2369	3	US-09-204-764-4	Sequence 4, Appl
C 32	14.2	71.0	2760	2	US-08-743-637B-20	Sequence 20, Appl
C 33	14.2	71.0	2760	3	US-08-526-840B-20	Sequence 20, Appl
C 34	14.2	71.0	2760	3	US-08-748-170A-3	Sequence 3, Appl
C 35	14.2	71.0	2760	3	US-09-047-148-1	Sequence 1, Appl
C 36	14.2	71.0	2978	4	US-09-293-395-1	Sequence 1, Appl
C 37	14.2	71.0	2978	4	US-09-668-648-1	Sequence 1, Appl
C 38	14.2	71.0	3451	2	US-08-743-637B-18	Sequence 1, Appl
C 39	14.2	71.0	3451	3	US-08-526-840B-18	Sequence 18, Appl
C 40	14.2	71.0	3509	2	US-08-327-832-4	Sequence 4, Appl
C 41	14.2	71.0	3509	2	US-08-828-584-4	Sequence 4, Appl
C 42	13.8	69.0	130	4	US-09-060-756-431	Sequence 431, App
C 43	13.8	69.0	1453	2	US-08-614-686A-2	Sequence 2, Appl
C 44	13.8	69.0	2131	1	US-08-325-071-64	Sequence 64, Appl
C 45	13.8	69.0	2131	4	US-08-461-004A-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-08-815-667-11
Sequence 11, Application US/08815667
Patent No. 5981173
GENERAL INFORMATION:
APPLICANT: Orth, Gerard
APPLICANT: Beaudenon, Sylvie
TITLE OF INVENTION: Longuet, Michele
TITLE OF INVENTION: TWO NOVEL GENITAL PAPILLOMAVIRUS (HPV)
TITLE OF INVENTION: TYPES, HPV68 AND HPV70, RELATED TO THE POTENTIALLY
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: FINNEGAN, HENDERSON, FARROW, GARRETT &
ADDRESSER: DUNNER, LLP
STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/815,667
FILING DATE: 11-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/011,650
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,458
FILING DATE: 15-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495-0153-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-815-667-11

Query Match 100.0%; Score 20; DB 2; Length 1059;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gaaacgcatatcctctgt 20
|||||
Db 782 GAAACGCCATACCTGCTGT 801

RESULT 2

US-08-815-667-10
; Sequence 10, Application US/08815667
; Patent No. 5981173
; GENERAL INFORMATION:
; APPLICANT: Orth, Gerard
; APPLICANT: Beaudenon, Sylvie
; APPLICANT: Longuet, Michele
; TITLE OF INVENTION: TWO NOVEL GENITAL PAPILLOMAVIRUS (HPV)
; TITLE OF INVENTION: TYPES, HPV68 AND HPV70, RELATED TO THE POTENTIALLY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; ADDRESSEE: DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/815,667
; FILING DATE: 11-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/011,650
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,458
; FILING DATE: 15-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495-0153-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1521 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-815-667-10

Query Match 100.0%; Score 20; DB 2; Length 1521;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gaaacgcatatcctctgt 20
|||||
Db 1045 GAAACGCCATACCTGCTGT 1064

RESULT 3

US-08-815-667-14
; Sequence 14, Application US/08815667
; Patent No. 5981173
; GENERAL INFORMATION:
; APPLICANT: Orth, Gerard
; APPLICANT: Beaudenon, Sylvie
; APPLICANT: Longuet, Michele
; TITLE OF INVENTION: TWO NOVEL GENITAL PAPILLOMAVIRUS (HPV)
; TITLE OF INVENTION: TYPES, HPV68 AND HPV70, RELATED TO THE POTENTIALLY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; ADDRESSEE: DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/815,667
; FILING DATE: 11-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/011,650
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,458
; FILING DATE: 15-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495-0153-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-815-667-14

Query Match 100.0%; Score 20; DB 2; Length 3283;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gaaacgcatatcctctgt 20
|||||
Db 1070 GAAACGCCATACCTGCTGT 1089

RESULT 4

US-08-474-542A-279
; Sequence 279, Application US/08474542A
; Patent No. 5527898
; GENERAL INFORMATION:
; APPLICANT: Bauer, Heidi M.
; APPLICANT: Gravit, Patti E.
; APPLICANT: Greer, Catherine E.
; APPLICANT: Imprim, Chaka C.
; APPLICANT: Manos, M. Michele
; APPLICANT: Resnick, Robert M.
; TITLE OF INVENTION: Detection of Human Papillomavirus by the
; TITLE OF INVENTION: Polymerase Chain Reaction

NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,542A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9234
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 279:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-542A-279

Query Match 84.0%; Score 16.8; DB 1; Length 394;
Best Local Similarity 90.0%; Pred. No. 4.5;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaaacggccatcactcgtct 20
||||| ||| ||| ||| ||| |||
Db 41 GAAACAGCCGCTACTGCTGT 60

RESULT 5
US-08-457-648-279
Sequence 279, Application US/08457648
Patent No. 5639871
GENERAL INFORMATION:
APPLICANT: Bauer, Heidi M.
APPLICANT: Gravitt, Patti E.
APPLICANT: Greer, Catherine E.
APPLICANT: Imprim, Chaka C.
APPLICANT: Manos, M. Michele
APPLICANT: Resnick, Robert M.
TITLE OF INVENTION: Detection of Human Papillomavirus by the
NUMBER OF SEQUENCES: 298
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,648
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Petry, Douglas A.
REGISTRATION NUMBER: 35,321
REFERENCE/DOCKET NUMBER: 9205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2974
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 279:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-457-648-279

Query Match 84.0%; Score 16.8; DB 1; Length 394;
Best Local Similarity 90.0%; Pred. No. 4.5;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaaacggccatcactcgtct 20
||||| ||| ||| ||| ||| |||
Db 41 GAAACAGCCGCTACTGCTGT 60

RESULT 6
US-09-011-745-7
Sequence 7, Application US/09011745
Patent No. 6165715
GENERAL INFORMATION:
APPLICANT: Collins, Mary KL
APPLICANT: Weiss, Robin A.
APPLICANT: Takeuchi, Yasuhiro
APPLICANT: Cossel, Francois-Lolc
TITLE OF INVENTION: Expression systems
FILE REFERENCE: 09/011,745
CURRENT APPLICATION NUMBER: US/09/011,745
CURRENT FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: PCT/GB96/02061
EARLIER FILING DATE: 1996-08-23
EARLIER APPLICATION NUMBER: GB9517263.1
EARLIER FILING DATE: 1995-08-23
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 6312
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Portion of
OTHER INFORMATION: construct
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4058)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4059)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4060)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4061)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4246)
OTHER INFORMATION: n is any nucleotide
FEATURE:

```
? NAME/KEY: misc_feature
? LOCATION: (4247)
? OTHER INFORMATION: n is any nucleotide
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (4248)
? OTHER INFORMATION: n is any nucleotide
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (4249)
? OTHER INFORMATION: n is any nucleotide
US-09-011-745-7
```

```
Query Match 76.0%; Score 15.2; DB 4; Length 6312;
Best Local Similarity 85.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 gaaagcgccatcctgctgt 20
|||||
DB 1496 gaaagcgccatcctgctgt 1515
```

```
RESULT 7
US-08-716-351A-1
; Sequence 1, Application US/08716351A
; Patent No. 6033905
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based
; TITLE OF INVENTION: Retroviral Vectors
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,351A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03784
; FILING DATE: 06-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-128-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8535 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..8535
; OTHER INFORMATION: /standard_name="GALV SEATO Genome"
US-08-716-351A-1
```

```
Query Match 76.0%; Score 15.2; DB 3; Length 8535;
Best Local Similarity 85.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 gaaagcgccatcctgctgt 20
|||||
DB 5872 gaaagcgccatcctgctgt 5891
```

```
RESULT 8
US-08-980-832-1
; Sequence 1, Application US/08980832B
; Patent No. 6291204
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsygankov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/08/980,832B
; CURRENT FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8625
; TYPE: DNA
; ORGANISM: Flavobacterium sp. R1534
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (8348)..(8349)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (8539)..(8540)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (8581)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (8590)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (8592)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (8602)..(8604)
US-08-980-832-1
```

```
Query Match 76.0%; Score 15.2; DB 4; Length 8625;
Best Local Similarity 85.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 gaaagcgccatcctgctgt 20
|||||
DB 2401 gaaagcgccatcctgctgt 2420
```

```
RESULT 9
US-08-980-832-27
; Sequence 27, Application US/08980832B
; Patent No. 6291204
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsygankov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/08/980,832B
; CURRENT FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 11233
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Plasmid pZee4
US-08-980-832-27
```

```
Query Match 76.0%; Score 15.2; DB 4; Length 11233;
Best Local Similarity 85.0%; Pred. No. 50;
```

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gaacggccatacctgtgt 20
|||||
Db 3083 gaacggccacgcctgtgt 3102

RESULT 10

US-08-535-276-1
; Sequence 1, Application US/08535276
; Patent No. 5977437
; GENERAL INFORMATION:
; APPLICANT: Villand, Per
; APPLICANT: Kleczkowski, Leszek
; APPLICANT: Olsen, Odd-Arne
; APPLICANT: Poulsen, Peter
; APPLICANT: Okkeles, Finn
; APPLICANT: Marcussen, Jan
; TITLE OF INVENTION: TRANSGENIC ORGANISM
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535,276
; FILING DATE: 05-FEB-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/01082
; FILING DATE: 07-APR-1994
; APPLICATION NUMBER: UK9307408.6
; FILING DATE: 08-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: DYOUL001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:

INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2037 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...2037
; OTHER INFORMATION: cDNA encoding large subunit
; OF ADP-glucose pyrophosphorylase (bep110)
US-08-535-276-1

Query Match 74.0%; Score 14.8; DB 2; Length 2037;
Best Local Similarity 88.9%; Pred. No. 63;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 aacggccatacctgtgt 20
|||
Db 530 AAGGGCCACACCTGCTGT 547

RESULT 11

US-08-449-644-4
; Sequence 4, Application US/08449644
; Patent No. 5856162
; GENERAL INFORMATION:
; APPLICANT: Schlössinger, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,644
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,244
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4651 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4317
US-08-449-644-4

Query Match 72.0%; Score 14.4; DB 2; Length 4651;
Best Local Similarity 93.8%; Pred. No. 1.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gaacggccatacctg 16
|||||
Db 3436 GAAAGTCCCATACCTG 3451

RESULT 12

US-08-087-244A-4
; Sequence 4, Application US/08087244A
; Patent No. 5863755
; GENERAL INFORMATION:
; APPLICANT: Schlössinger, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang

```

; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,244A
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4651 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4317
; US-08-087-244A-4

```

```

Query Match 72.0%; Score 14.4; DB 2; Length 4651;
Best Local Similarity 93.8%; Pred. No. 1.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 gaacggccatactg 16
      ||||| |||||
Db 3436 GAAACTGCCATACCTG 3451

```

```

RESULT 13
US-08-961-810-17/c
; Sequence 17, Application US/08961810
; Patent No. 6165713
; GENERAL INFORMATION:
; APPLICANT: Liskay, Robert M.
; APPLICANT: Bronner, C. Eric
; APPLICANT: Baker, Sean M.
; APPLICANT: Bollag, Roni J.
; APPLICANT: Kolodner, Richard D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &
; ADDRESSEE: Heuser
; STREET: 520 S.W. Yamhill Street, Suite 200
; CITY: Portland
; STATE: Oregon
; COUNTRY: U.S.A.
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,810
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Rysselberghe, Pierre C.
; REGISTRATION NUMBER: 33,557
; REFERENCE/DOCKET NUMBER: OHSU 306B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 224-6655
; TELEFAX: (503) 295-6679
; TELEX: 360619
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-961-810-17

Query Match 71.0%; Score 14.2; DB 4; Length 618;
Best Local Similarity 84.2%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 aaacggccatactgctgt 20
      ||| ||||| ||||| ||
Db 521 AAAAGGCCATACCTGGGCT 503

RESULT 14
US-08-352-902D-17/c
; Sequence 17, Application US/08352902D
; Patent No. 6191268
; GENERAL INFORMATION:
; APPLICANT: Liskay, Robert M.
; APPLICANT: Bronner, C. Eric
; APPLICANT: Baker, Sean M.
; APPLICANT: Bollag, Roni J.
; APPLICANT: Kolodner, Richard D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &
; ADDRESSEE: Heuser
; STREET: 520 S.W. Yamhill Street, Suite 200
; CITY: Portland
; STATE: Oregon
; COUNTRY: U.S.A.
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/352,902D
; FILING DATE: 09-Dec-1994
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Rysselberghe, Pierre C.
; REGISTRATION NUMBER: 33,557
; REFERENCE/DOCKET NUMBER: OHSU 306B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 224-6655
; TELEFAX: (503) 295-6679
; TELEX: 360619
; INFORMATION FOR SEQ ID NO: 17:

```


;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-08-352-902D-17

Query Match 71.0%; Score 14.2; DB 4; Length 618;
Best Local Similarity 84.2%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 aaacggccataccctgctgt 20
|||||
DB 521 AAAAGGCCATACCTGGGGT 503

RESULT 15
US-09-293-395-6/C
; Sequence 6, Application US/09293395
; Patent No. 6174860
; GENERAL INFORMATION:
; APPLICANT: Kramer, Vance
; APPLICANT: Morgan, Michael
; APPLICANT: Anderson, Arne
; TITLE OF INVENTION: NOVEL INSECTICIDAL TOXINS
; FILE REFERENCE: PB/5-30472A/USN
; CURRENT APPLICATION NUMBER: US/09/293,395
; CURRENT FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1290
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1287)
; OTHER INFORMATION: JHE-likeorf2 of pCIB0381
US-09-293-395-6

Query Match 71.0%; Score 14.2; DB 4; Length 1290;
Best Local Similarity 84.2%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gaacggccataccctgctg 19
|||||
DB 1259 GAAATGCAATCCCTGCTG 1241

Search completed: April 3, 2002, 10:31:12
Job time: 3463 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

-Run on: April 3, 2002, 10:28:35 ; Search time 3285.41 Seconds
(without alignments)
65.415 Million cell updates/sec

Title: US-09-885-799-488

Perfect score: 20

Sequence: 1 gaacagccatactgctgt 20

Scoring table: IDENTITY_NUC

Gap 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estfun:*

2: em_esthum:*

3: em_estin:*

4: em_estom:*

5: em_estpl:*

6: em_estba:*

7: em_estro:*

8: em_estov:*

9: em_hic:*

10: gb_est1:*

11: gb_est2:*

12: gb_hic:*

13: gb_gss:*

14: em_gss_fun:*

15: em_gss_hum:*

16: em_gss_inv:*

17: em_gss_pln:*

18: em_gss_pro:*

19: em_gss_rod:*

20: em_gss_vrt:*

21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	length	DB ID	Description
1	18	90.0	522	AZ156803	AZ156803 SP_0014_A
2	17.4	87.0	469	AQ440154	AQ440154 HS_5078_A
3	17.4	87.0	589	11 BG590213	BG590213 ES7498043
4	16.8	84.0	390	11 BG321873	BG321873 ib30c04.y
5	16.8	84.0	917	13 CNS075GK	AL430074 clone BAO
6	16.4	82.0	166	13 AZ122035	AZ122035 RPI-23-1
7	16.4	82.0	378	13 AZ122006	AZ122006 RPI-23-1
8	16.4	82.0	775	13 CNS01XDP	AL171574 Tetraodon
9	16.4	82.0	831	13 CNS04SHW	AL305213 Tetraodon
10	15.8	79.0	377	13 AQ116621	AQ116621 HS_3076_A
11	15.8	79.0	399	13 AQ124659	AQ124659 HS_2267_B
12	15.8	79.0	505	10 AA77625	AA77625 z195a12.s

C 13	15.8	79.0	523	13 TA12C07P	AL452132 T. brucei
C 14	15.8	79.0	541	10 BE462403	BE462403 EST324668
C 15	15.8	79.0	541	10 BE443097	BE443097 WHE1113_F
C 16	15.8	79.0	556	10 BE443075	BE443075 WHE1113_H
C 17	15.8	79.0	563	10 AW648257	AW648257 EST326711
C 18	15.8	79.0	573	10 AW651346	AW651346 EST329800
C 19	15.8	79.0	618	10 AW648252	AW648252 EST326706
C 20	15.8	79.0	635	13 AZ359308	AZ359308 LM0102P01
C 21	15.8	79.0	643	10 AW041697	AW041697 EST284561
C 22	15.8	79.0	647	13 AO952751	AO952751 Sheared D
C 23	15.8	79.0	652	13 AO648711	AO648711 RPI93-DP
C 24	15.8	79.0	691	10 AV399761	AV399761 AV399761
C 25	15.8	79.0	732	10 A1778708	A1778708 EST259587
C 26	15.8	79.0	738	11 BG124041	BG124041 EST469687
C 27	15.8	79.0	789	11 BG124254	BG124254 EST469900
C 28	15.8	79.0	809	10 A1743718	A1743718 wg53a11.X
C 29	15.8	79.0	832	13 AZ126288	AZ126288 OSJNB007
C 30	15.8	79.0	886	13 CNS02WFE	AL216995 Tetraodon
C 31	15.8	79.0	894	11 BF206161	BF206161 601869337
C 32	15.8	79.0	934	13 CNS030TV	AL222700 Tetraodon
C 33	15.8	79.0	1074	13 CNS05GRJ	AL336664 Tetraodon
C 34	15.6	78.0	868	13 CNS03DK2	AL239228 Tetraodon
C 35	15.4	77.0	220	10 BB026444	BB026444 BR026444
C 36	15.4	77.0	321	11 Z37666	Z37666 ATSA131.Ye
C 37	15.4	77.0	340	10 AV559880	AV559880 AV559880
C 38	15.4	77.0	387	13 B32964	B32964 HS-1016-A2-
C 39	15.4	77.0	404	10 AW119307	AW119307 M19 Neosp
C 40	15.4	77.0	406	11 BG447017	BG447017 GA_EB004
C 41	15.4	77.0	444	13 AQ886328	AQ886328 HS_5542_B
C 42	15.4	77.0	552	13 AO677032	AO677032 HS_2148_B
C 43	15.4	77.0	555	10 AV720800	AV720800 AV720800
C 44	15.4	77.0	566	13 BH014939	BH014939 TDGRS69TH
C 45	15.4	77.0	577	11 BF330972	BF330972 PM3-BT034

ALIGNMENTS

RESULT	1				
LOCUS	AZ156803	522 bp	DNA	29-AUG-2000	
DEFINITION	SP_0014.AL_R12_T7 Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plate-14 Col=23 Row=C, DNA sequence.				
ACCESSION	AZ156803				
VERSION	AZ156803.1	GI:8309402			
KEYWORDS	GSS.				
SOURCE	Strongylocentrotus purpuratus.				
ORGANISM	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae; Strongylocentrotus.				
REFERENCE	1 (bases 1 to 522)				
AUTHORS	Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R., Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and Hood,L.				
TITLE	A sea urchin genome project: Sequence scan, virtual map, and additional resources				
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 97 (17), 9514-9518 (2000)				
MEDLINE	20402566				
COMMENT	Contact: Cameron, RA, Davidson, EH, Hood, L Division of Biology 156-29 California Institute of Technology Pasadena California 91125, USA Tel: (626) 395-8421 Fax: (626) 793-3047 Email: acameron@caltech.edu Plate: 14 row: C column: 23 Seq primer: T7 Class: BAC ends High quality sequence stop: 522. Location/Qualifiers				

```

Query Match      87.0%; Score 17.4; DB 13; Length 469;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaacggccatacctgctgt 20
||||||| ||||||| |||
Db 234 GAACGGCAATACCTGCTGTT 253

RESULT 3
BG590213/C
LOCUS BG590213 589 bp mRNA EST 12-APR-2001
DEFINITION EST438043 P. infestans-challenged leaf Solanum tuberosum cDNA clone
BPL14G4 5' sequence, mRNA sequence.
ACCESSION BG590213
VERSION BG590213.1 GI:13608341
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 589)
Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukoianov, A.,
Rangel, P., Haberlach, G.T., Cho, J., Chiemingo, A., Bougri, O., Buehl
, C.R., Ronning, C.M., Helgeson, J., and Baker, B.
Generation of ESTs from Potato Leaves Challenged with Phytophthora
infestans, Incompatible Reaction
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-771-6195, email cdm@resgen.com
Seq primer: M13F-R.

FEATURES
source
1..589
Location/Qualifiers
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone_lib="P. infestans-challenged leaf"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; whole plants were challenged with 450,000
sporangia/ml P. infestans US-1(US 940501) in Biotron
(Madison, WI). Leaf tissue was collected at 1, 2,
5, 12, and 24 hours post-challenge and frozen in liquid
nitrogen immediately upon removal. Kennebec plants showed
no signs of HR. Katakadin plants (susceptible to P.
infestans US-1) were used as controls and showed
infection. NOTE: We cannot exclude the possibility that
this sequence is actually derived from Phytophthora rather
than potato."

BASE COUNT 164 a 114 c 142 g 169 t
ORIGIN

Query Match      87.0%; Score 17.4; DB 11; Length 589;
Best Local Similarity 94.7%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 aaacggccatacctgctgt 20
||||||| ||||||| |||||||
Db 182 AAACGGCCATTCCTGCTGT 164

RESULT 4
BG321873
LOCUS BG321873 390 bp mRNA EST 05-JUL-2001
DEFINITION ib30c04.y1 Melton Mouse Newborn Pancreas Mus musculus cDNA 5'

```

similar to SW:POL.SMSAY P0359 POL POLYPROTEIN [CONTAINS: REVERSE
TRANSCRIPTASE ; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BC321873.1 GI:13151551

EST.

house mouse.

Mus musculus

Eukaryota;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 390)

REFERENCE

AUTHORS

Meitner, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,

Schmitt, A., Theisinger, B., Ritter, E., Ronko, I., Bennett, J., Cardenas

, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.

, Jackson, Y., and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Juliana Brown

(brown@fas.harvard.edu)

Seq primer: -40RP from Gibco.

Location/Qualifiers

1..390

/organism="Mus musculus"

/strain="ICR"

/db_xref="taxon:10090"

/clone_lib="Melton Mouse Newborn Pancreas"

/sex="Both"

/tissue_type="Total Pancreas"

/dev_stage="Day 1 through day 4 postnatal, mixed"

/lab_host="TOP10"

/note="Organ: Pancreas; Vector: pZeo-2; Site.1: Not 1;

Site.2: Xho I; Library constructed using SuperScript

Plasmid Library kit (Life Technologies). cDNA made by

oligo-dT priming. Xho I site destroyed during cloning.

Size selected by column fractionation. Primary library,

unamplified."

BASE COUNT

ORIGIN

102 a 109 c 95 g 84 t

Query Match

Best Local Similarity

Matches

18; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

QY

1 gaaacgcccatacctgctgt 20

||||| ||||| ||||| |||||

Db

87 GAAAGGCCCATACCTGGTGT 106

RESULT

5

CNS075GK/c

LOCUS

CNS075GK 917 bp DNA GSS

DEFINITION

clone BA0AB035E11 of library BA0AB from strain CluB 210 of

Kluyveromyces lactis, genomic survey sequence.

ACCESSION

AL430074

VERSION

AL430074.1 GI:12213268

KEYWORDS

GSS.

SOURCE

Kluyveromyces lactis.

ORGANISM

Kluyveromyces lactis

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Kluyveromyces.

REFERENCE

1 (bases 1 to 917)

AUTHORS

Bolotin-Fukuhara, M., Toffano-Niccho, C., Artiguenave, F.,

Duchateau-Nguyen, G., Lemaire, P., Marmelisse, R., Montrocher, R.,

Robert, C., Terrier, M., Wincker, P., and Wesolowski-Louvel, M.

Genomic exploration of the hemiascomycetous yeasts: 11.

Kluyveromyces lactis

FEBS Lett. 487 (1), 66-70 (2000)

20584721

2 (bases 1 to 917)

Souchet, J.B., Aglie, M., Artiguenave, F., Blandin, G.,

Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,

de-Montigny, J., Dujon, B., Durieux, P., Lepingle, A., Llorente, B.,

Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S.,

Saurin, W., Tekala, F., Toffano-Niccho, C., Wesolowski-Louvel, M.,

Wincker, P., and Weissenbach, J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of

yeast species for molecular evolution studies

FEBS Lett. 487 (1), 3-12 (2000)

20584711

3 (bases 1 to 917)

Genoscope.

Direct Submission

Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,

2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :

seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

This GSS is part of a random genomic sequencing program of thirteen

yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces

exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,

Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces

lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia

angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,

Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to

5 kb were prepared and both extremities were sequenced. See

keywords for description of this sequence and for the sequence of

the other extremity of this insert.

FEATURES

source

1..917

/location/Qualifiers

/organism="Kluyveromyces lactis"

/strain="CLIB 210"

/variety="lactis"

/db_xref="taxon:28985"

/clone="BA0AB035E11"

/clone_lib="BA0AB"

<2..>835

/note="similar to Saccharomyces cerevisiae ORF YDR293c [

SSD1; involved in the tolerance to high concentration of

Caz+]"

/evidence=not_experimental

BASE COUNT

ORIGIN

305 a 198 c 190 g 224 t

Query Match

Best Local Similarity

Matches

18; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

QY

1 gaaacgcccatacctgctgt 20

||||| ||||| ||||| |||||

Db

296 GAAATGGGCATACCTGCTGT 277

RESULT

6

AZ122035/c

LOCUS

AZ122035 166 bp DNA GSS

DEFINITION

RPC1-23-11N1.TV RPC1-23 Mus musculus genomic clone RPC1-23-11N1,

DNA sequence.

ACCESSION

AZ122035

VERSION

AZ122035.1 GI:7789536

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 166)

JOURNAL

Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

Location/Qualifiers
1..775

/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="202C22"
/clone_lib="G"

BASE COUNT 172 a 186 c 227 g 189 t 1 others
ORIGIN

Query Match 82.0% Score 16.4; DB 13; Length 775;
Best Local Similarity 94.4% Pred. No. 7.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 gaacggccatctctgt 18
|||||
Db 591 GAACGGCCAACTCTCT 608

RESULT 9 831 bp DNA GSS
CNS04SHW/LOCUS
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Or1 end of clone
007D07 of library H from Tetraodon nigroviridis, genomic survey
sequence.
AL305213.1 GI:8196344
GSS: genome survey sequence.
Tetraodon nigroviridis.
Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 831)
Roest-Crolius,H., Jallion,O., Dasilva,C., Filames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 831)
Roest-Crolius,H., Jallion,O., Dasilva,C., Bonneau,L., Fisher,C.,
Bernot,A., Filames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 831)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

BASE COUNT 253 a 176 c 209 g 192 t 1 others
ORIGIN

Query Match 82.0% Score 16.4; DB 13; Length 831;
Best Local Similarity 94.4% Pred. No. 7.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 aaacggccatctctgt 19
|||||
Db 664 ACACGGCCATCTCTCT 647

RESULT 10 377 bp DNA GSS
A0116621/LOCUS
DEFINITION HS_3076_AL.D02_MR.CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-3076 Col-3 Row-G, DNA sequence.
A0116621
VERSION A0116621.1 GI:3493832
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 377)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3076 row: G column: 3
Class: BAC ends
High quality sequence stop: 377.
Location/Qualifiers
1..377
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3076 Col=3 Row=G"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT 94 a 89 c 106 g 86 t 2 others
ORIGIN

Query Match 79.0% Score 15.8; DB 13; Length 377;
Best Local Similarity 89.5% Pred. No. 1.4e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 gaacggccatctctgt 19
|||||
Db 255 GAACGGCCACACCGCTG 237

RESULT 11 399 bp DNA GSS
A0124659/LOCUS
DEFINITION HS_2267_B1.H06_MF.CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-2267 Col-11 Row-P, DNA sequence.
A0124659
VERSION A0124659.1 GI:3501825
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 399)
AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
COMMENT	99380389 Contact: Mahairas GG, Wallace JC, Hood L

Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 267 row: P column: 11
 Class: BAC ends
 High quality sequence stop: 399.
 Location: 0.0001515

BASE COUNT	ORIGIN
111 a	98 c 84 g 105 t 1 others

Query Match	79.0%	Score 15.8	DB 13	Length 399
Best Local Similarity	89.5%	Pred No. 1.4e+03		
Matches 17	Conservative 0	Mismatches 2	Indels 0	Gaps 0
OY	2	aaacggccatactctgcgtgt	20	
Db	232	AAACGACCACTACTGCTGT	250	

RESULT	12
AA777625	
LOCUS	AA777625
DEFINITION	505 bp mRNA
VERSION	EST
ACCESSION	U15412.S1 Soares_tet1_liver_spleen_INTS.S1
VERSION	AA777625
KEYWORDS	AA777625.1 GI:2837104
SOURCE	EST.
ORGANISM	human.
	Homo sapiens
	05-FEB-1998

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 505)	Mammalia; Eutheria; Primates; Carnivora; Hominiidae; Homo.			
	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucabada, T., Lacy, M., Le, N., Lennarz, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., Wille, T., Wylie, T., Waterson, R. and Wilson, R.	WashU-NCI human EST project	Unpublished (1997)	Contact: Wilson RK

FEATURES

Email: esi@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence strip: 467.
Location/Qualifiers

Source

```

SOURCE
1. 505
/organism="Homo sapiens"
/db_xref="GDB:1352767"
/db_xref="taxon:9606"
/clone="IMAGE:448510"
/clone_lib="Soares_fetal_liver_spleen_INTLS_S1"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: liver and spleen; Vector: pT73D (pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INTLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAGAGATTATTAAGCATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaudo."
BASE COUNT
169 a 80 c 82 g 174 t
ORIGIN

```

```

Query Match: 79.0% Score 15.8; DB 10; Length 505;
Best Local Similarity 89.5%; Pred No. 1.4e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 aaagcgcataccctgcgtct 20
    | | | | | | | | | | | | | |
Db 423 ATRCAGCCATACGTCGT 441

```

RESULT	13
LOCUS	TA12C07P/c
DEFINITION	TA12C07P 523 bp DNA GSS 13-DEC-2000 T. brucei sheared genomic DNA clone 12c07, forward sequence.
ACCESSION	AL452132
VERSION	AL452132.1 GI:11833426
KEYWORDS	GSS.
SOURCE	Trypanosoma brucei.
ORGANISM	Trypanosoma brucei

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 523)	Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.	Direct Submission	Submitted (10-DEC-2000)	Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA. E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk

Rocky, M. Genomic DNA isolated from a cloned population of *Trypanosoma brucei* (TREP927/4 Ghat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The *vt* 1 method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barrell, Oxford University Press, 1999). Email: nelsayed@telir.org

Details of *T. brucei* sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/projects/T_brucei/.

FEATURES	Location/Qualifiers
SOURCE	1. 523
	/organism="Trypanosoma brucei"
	/strain="TREU927"
	/db_xref="taxon:5691"
	/clone="12c07"
BASE COUNT	177 a 140 c 125 g 81 t

ORIGIN

Query Match 79.0%; Score 15.8; DB 13; Length 523;
 Best Local Similarity 89.5%; Pred. No. 1.5e+03;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gaaagcgcatacctgctg 19
 ||||| ||||| ||||| |||||
 DB 358 GAACAGCCATTCCTGCTG 340

RESULT 14
 BE462403 541 bp mRNA EST 18-MAY-2001
 LOCUS EST324668 tomato flower buds 0-3 mm, Cornell University
 DEFINITION Lycopersicon esculentum cDNA clone cT0A12J22, mRNA sequence.
 BE462403
 VERSION BE462403.1 GI:9508172
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 541)
 van der Hoeven, R.S., Bezzeredes, J.L., Matern, A.L., Holt, T.E., Liang
 F., Hansen, T., Craven, M.B., Bowman, C.L., Romling, C.M., Nierman, W.,
 Fraser, C.M., Martin, G.B., Giovannoni, J.J., and Tanksley, S.D.
 Generation of ESTs from tomato flower tissue, 0-3 mm buds
 Unpublished (1999)
 CONTACT: CUGI
 CLEMSON UNIVERSITY GENOMICS INSTITUTE
 CLEMSON UNIVERSITY
 100 JORDAN HALL, CLEMSON, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 Location/Qualifiers
 1..541
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cT0A12J22"
 /clone_1lb="tomato flower buds 0-3 mm, Cornell University"
 /tissue_type="flower"
 /dev_stage="0-3mm buds"
 /note="Vector: Bluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; supplier: Tanksley; Flower buds and flowers were
 taken from greenhouse plants (4-8 wks old; TA496). They
 were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."

BASE COUNT 147 a 113 c 115 g 166 t

ORIGIN

Query Match 79.0%; Score 15.8; DB 10; Length 541;
 Best Local Similarity 89.5%; Pred. No. 1.5e+03;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 aaarqccatacctgctg 20
 ||||| ||||| ||||| |||||
 DB 501 AAACAGCCATTCCTGCTG 483

RESULT 15
 BE443097 544 bp mRNA EST 25-JUL-2000
 LOCUS MHE1113.F10.K192S wheat etiolated seedling root normalized cDNA
 DEFINITION library Triticum aestivum cDNA clone MHE113.F10.K19, mRNA
 sequence.
 BE443097
 VERSION BE443097.1 GI:9442625

KEYWORDS EST.
 SOURCE bread wheat.
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticaceae; Triticum.
 1 (bases 1 to 544)
 Rausch, C.J., Seaton, C.L., Tong, J.C., and Zhang, D.
 The structure and function of the expressed portion of the wheat
 genomes - Normalized root cDNA library
 unpublished (2000)
 CONTACT: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105959773
 Fax: 5105959818
 Email: Oanderson@pw.usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: Stratagene SK primer.

FEATURES
 Location/Qualifiers
 1..544
 /organism="Triticum aestivum"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="MHE113.F10.K19"
 /clone_1lb="wheat etiolated seedling root normalized cDNA
 library"
 /tissue_type="Root"
 /dev_stage="Five day old etiolated seedling"
 /lab_host="E. coli DH10B"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid
 pluscript SK; Site_1: EcoRI; Site_2: XhoI; Seeds were
 surface-sterilized, germinated and grown aseptically in
 the dark at room temperature on filter paper with water,
 nystatin and cefotaxime in covered crystalization
 dishes. Roots were harvested. The tissue, total RNA, and
 poly(A) RNA were prepared, a cDNA library was made in the
 TI Close lab (Choi, Close, Fenton) at the University of
 California, Riverside. The cDNA clones were in vivo
 excised to give pluscript phagemids before
 normalization was carried out. The mass excision of
 phagemid library and normalization were done in HT Nguyen
 lab by D. Zhang at Texas Tech University. Normalization
 protocol used was that of Soares. Plasmid DNA
 preparations and DNA sequencing were performed in the UD
 Anderson lab (all other authors)."

BASE COUNT 92 a 177 c 178 g 96 t 1 others

ORIGIN

Query Match 79.0%; Score 15.8; DB 10; Length 544;
 Best Local Similarity 89.5%; Pred. No. 1.5e+03;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gaaagcgcatacctgctg 19
 ||||| ||||| ||||| |||||
 DB 98 GAACAGCCATTCCTGCTG 116

Search completed: April 3, 2002, 10:28:38
 Job time: 3309 sec

Thu Apr 4 09:39:32 2002

us-09-885-799-488.rst

Myers
09/885799

09/885799

L1 FILE 'REGISTRY' ENTERED AT 10:21:20 ON 04 APR 2002
27 SEA ABB=ON PLU=ON ATGCACTGAAGTAACTAAGG|CACTGAAGTAACTAAG
GAAG|GAAACGGCCATACCTGCTGT|CGGCCATACCTGCTGTATATAG/SQSN

L2 FILE 'CAPLUS' ENTERED AT 10:26:28 ON 04 APR 2002
14 SEA ABB=ON PLU=ON L1

L2 ANSWER 1 OF 14 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:693572 CAPLUS
DOCUMENT NUMBER: 135:268104
TITLE: Genotyping kit for diagnosis of human
papillomavirus infection
INVENTOR(S): Park, Tae-Shin; Kim, Sung-Keun; Kim, Jin-Hee;
Park, Mi-Sun
PATENT ASSIGNEE(S): Biomedlab Corporation, S. Korea
SOURCE: PCT Int. Appl., 50 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001068915	A1	20010920	WO 2000-KR1213	20001026
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				

PRIORITY APPLN. INFO.: KR 2000-13161 A 20000315
AB The present invention relates to a genotyping kit for diagnosing
patients infected with human papillomavirus (HPV), and a method for
diagnosis of HPV infection by genotyping specimen DNA isolated from
the patients using the said kit. The genotyping kit of the
invention comprises a DNA chip with probes that have nucleotide
sequences complementary to DNA of HPV, primers for amplifying the
DNA of a sample by PCR, and means for labeling hybridized sample DNA
with the DNA chip. The method for diagnosis of HPV infection
comprises the steps of amplifying DNA obtained from a sample using
the primers of the kit, applying the amplified DNA to the DNA chip
and hybridizing the amplified DNA and the probes of the DNA chip,
and detecting DNA bound on the surface of the DNA chip by labeling
hybridized DNA. In accordance with the invention, the genotyping
kit may be practically applied to the early diagnosis, prevention
and treatment of cervical cancer, since the kit can easily diagnose
HPV infection, and can exactly det. the genotype of the HPV.

IT 195231-89-1

RL: PRP (Properties)

(unclaimed nucleotide sequence; genotyping kit for diagnosis of
human papillomavirus infection)

REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR
THIS RECORD. ALL CITATIONS AVAILABLE IN
THE RE FORMAT

Searcher : Shears 308-4994

09/885799

L2 ANSWER 2 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2000:661694 CAPLUS

DOCUMENT NUMBER: 134:217713

TITLE: Amplification with molecular beacon primers and reverse line blotting for the detection and typing of human papillomaviruses

AUTHOR(S): Jordens, J. Z.; Lanham, S.; Pickett, M. A.;

CORPORATE SOURCE: Amarasekara, S.; Abeywickrema, I.; Watt, P. J. Department of Molecular Microbiology, University of Southampton, Southampton General Hospital, Southampton, SO16 6YD, UK

SOURCE: Journal of Virological Methods (2000), 89(1-2), 29-37

CODEN: JVMEDH; ISSN: 0166-0934

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB A novel method for the detection and typing of human papillomavirus (HPV) was developed using mol. beacon primers. The method is based on the use of HPV-specific primers contg. a hairpin loop structure in which fluorescent donor and quencher groups are held in close proximity such that fluorescence is quenched. Amplification of the target sequence results in the opening of the loop and the resulting fluorescence can be detected on a sequence detector system (SDS) 7700 (Applied Biosystems), as used for TaqMan assays. Fluorescent amplicons were identified on the SDS 7700 and then typed by a single hybridization with specific probes immobilized in lines on a nylon membrane and detected on a fluorescent scanner. This novel beacon primer method compared well with conventional PCR for cervical scrape specimens. The combination of the beacon primer method and reverse line blotting should enable large-scale population studies of HPV infection.

IT 165309-07-9

RL: ARG (Analytical reagent use); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)

(amplification with mol. beacon primers and reverse line blotting for detection and typing of human papillomaviruses)

REFERENCE COUNT: 18 THERE ARE 18 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 3 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1999:718867 CAPLUS

DOCUMENT NUMBER: 131:347519

TITLE: DNA molecules encoding human papillomavirus 68

E6, E7, L1 proteins, their sequences and uses

INVENTOR(S): Orth, Gerard; Beaudenon, Sylvie; Longuet, Michele

PATENT ASSIGNEE(S): Institut Pasteur, Fr.; Institut Nationale de la Sante et de la Recherche Medicale

SOURCE: U.S., 25 pp., which

CODEN: USXXAM

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO. KIND DATE APPLICATION NO. DATE

Searcher : Shears 308-4994

09/885799

US 5981173	A	19991109	US 1997-815667	19970211
PRIORITY APPLN. INFO.:			US 1996-11650P	P 19960214

AB The invention provides human papillomavirus type 68 (HPV68) DNA mols. isolated from a low grade cervical intraepithelial neoplasia cell line encoding E6, E7, L1 and L2 proteins. The invention also provides a method for detection of HPV68 virus in a biol. sample using the HPV68 DNA mols. presented in the invention as probes. The invention further provides a method for recombinant prodn. of HPV68 DNA using a cloning vector contg. the HPV68 DNA sequences presented in the invention and transforming a host cell with said vector. The DNA sequences of HPV68 genes encoding E6, E7, and L1 proteins are provided. The invention also provided the DNA sequences of human papillomavirus type 70 (HPV70) E6, E7, L1 and L2 protein genes, and compared the sequences of HPV68 and HPV70 to human papillomavirus type 39 (HPV39) and ME180-HPV (cervical carcinoma-derived ME 180 cell line). The invention showed that HPV68 and HPV70 were detected in genital intraepithelial neoplasia from three patients and one patient, resp. Comparison with sequence data in the literature indicates that the subgenomic ME180-HPV DNA fragment, cloned from a carcinoma cell line, corresponds to a HPV68 subtype and that several HPV DNA fragments amplified by PCR from genital neoplasia represent worldwide distributed variants of HPV68 and HPV70.

IT 175702-95-1, GenBank U22461

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(nucleotide sequence; DNA mols. encoding human papillomavirus 68 E6, E7, L1 proteins, and comparison to human papillomavirus 70 sequences)

IT 250163-20-3 250163-21-4

RL: PRP (Properties)

(unclaimed nucleotide sequence; dNA mols. encoding human papillomavirus 68 E6, E7, L1 proteins, their sequences and uses)

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR
THIS RECORD. ALL CITATIONS AVAILABLE IN
THE RE FORMAT

L2 ANSWER 4 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1996:235613 CAPLUS

DOCUMENT NUMBER: 124:309077

TITLE: Human papillomavirus type 70 genome cloned from overlapping PCR products: Complete nucleotide sequence and genomic organization

AUTHOR(S) : Forslund, Ola; Hansson, Bengt Goeran

CORPORATE SOURCE: Department Medical Microbiology, Lund

University, Malmö, S-205 02, Swed.

SOURCE: J. Clin. Microbiol. (1996), 34(4), 802-9

CODEN: JCMIDW; ISSN: 0095-1137

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The genome of human papillomavirus (HPV) type 70 (HPV 70), isolated from a cervical condyloma, was obtained by cloning overlapping PCR products. By automated DNA sequence anal., the genome was found to consist of 7,905 bp with a G+C content of 40%. The genomic organization showed the characteristic features shared by other sequenced HPVs. Nucleotide sequence comparison with previously known HPV types demonstrated the closet homol. with HPV 68 (82%),

Searcher : Shears 308-4994

09/885799

HPV 39 (82%), HPV 18 (70%), HPV 45 (70%), and HPV 59 (70%). Comparison with seven other partially sequenced HPV 70 isolates showed homologies of between 100 and 99.5%. Cloning of overlapping PCR products and automated DNA sequence anal. was found to be a feasible method of obtaining full-length sequences of HPVs.

IT 164961-97-1, GenBank U21941

RL: PRP (Properties)

(nucleotide sequence; complete nucleotide sequence and genomic organization of human papillomavirus type 70 genome cloned from overlapping PCR products)

L2 ANSWER 5 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1996:227296 CAPLUS

DOCUMENT NUMBER: 124:280734

TITLE: Intratype variation in 12 human papillomavirus types: a worldwide perspective

AUTHOR(S): Stewart, Ann-Charlotte M.; Eriksson, Annika M.; Manos, M. Michele; Munoz, Nubia; Bosch, F. Xavier; Peto, Julian; Wheeler, Cosette M.

CORPORATE SOURCE: Dep. Cell Biol., Univ. New Mexico Sch. Med., Albuquerque, NM, 87131, USA

SOURCE: J. Virol. (1996), 70(5), 3127-36

CODEN: JOVIAM; ISSN: 0022-538X

DOCUMENT TYPE: Journal

LANGUAGE: English

AB In this study, we have examd. intratype human papillomavirus (HPV) sequence variation in a worldwide collection of cervical specimens. Twelve different HPV types including HPV-18, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68 (ME180), MM9/PAP238A (recently designated HPV-73), and a novel partial genomic HPV sequence designated MM4/W13B were analyzed in this study. Cervical specimens were collected as part of epidemiol. investigations conducted in New Mexico and an internal study of invasive cervical cancer (IBSCC). Specimens from several countries including Argentina, Brazil, Bolivia, Benin, Cuba, Colombia, Chile, Germany, Mali, Panama, Paraguay, Spain, Algeria, Uganda, Guinea, Tanzania, Indonesia, Philippines, Thailand, and the United States were evaluated. Specimen DNAs were subjected to amplification with the MY09/11 L1 consensus PCR system. The PCR products were cloned, and an approx. 410-bp region in the L1 open reading frame was sequenced from 146 specimens (.apprx.60,000 bp). Within a single HPV type, nucleotide diversity varied between 0.2 and 2.9% (i.e., between any pair of variants) and the majority of nucleotide changes were synonymous (amino acid conserving). These data provide information pertinent to HPV diagnostic probe development and are potentially relevant to future rational vaccine strategies. Similarly, amino acid diversity varied between 0 and 5.1%. Some of these amino acid changes may represent markers of intertype evolutionary relationships. Presuming that HPVs have evolved under the same constraints as their corresponding hosts, the limited genetic diversity obsd. for all HPVs studied to date may reflect an evolutionary bottleneck occurring in both virus and host populations.

IT 174054-08-1, GenBank U45924 174054-09-2, GenBank

U45925 174054-10-5, GenBank U45926 174054-11-6,

GenBank U45927 174054-12-7, GenBank U45928

RL: PRP (Properties)

(nucleotide sequence; world-wide intratype variation in 12 human

09/885799

papillomavirus types)

L2 ANSWER 6 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1996:190200 CAPLUS

DOCUMENT NUMBER: 124:252008

TITLE: Generation of entire human papillomavirus genomes by long PCR: frequency of errors produced during amplification. [Erratum to document cited in CA123:307451]

AUTHOR(S): Stewart, Charlotte M.; Gravitt, Patti E.; Cheng, Suzanne; Wheeler, Cosette M.

CORPORATE SOURCE: USA

SOURCE: Genome Res. (1996), 6(2), 166

CODEN: GEREFS

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The errors were not reflected in the abstr. or the index entries.

IT 160074-98-6P, GenBank U12476

RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)

(nucleotide sequence; generation of entire human papillomavirus genomes by long PCR and amplification of novel HPV genomes (Erratum))

L2 ANSWER 7 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1996:182462 CAPLUS

DOCUMENT NUMBER: 124:280570

TITLE: Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70, related to the potentially oncogenic HPV39

AUTHOR(S): Longuet, Michele; Beaudenon, Sylvie; Orth, Gerard

CORPORATE SOURCE: Unite Papillomavirus, Inst. Pasteur, Paris, 75724, Fr.

SOURCE: J. Clin. Microbiol. (1996), 34(3), 738-44

CODEN: JCMIDW; ISSN: 0095-1137

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The genomes of two novel human papillomavirus (HPV) types, HPV68 and HPV70, were cloned from a low-grade cervical intraepithelial neoplasia and a vulvar papilloma, resp., and partially sequenced. Both types are related to HPV39, a potentially oncogenic virus. HPV68 and HPV70 were also detected in genital intraepithelial neoplasia from three patients and one patient, resp. Comparison with sequence data in the literature indicates that the subgenomic ME180-HPV DNA fragment, cloned from a carcinoma cell line, corresponds to an HPV68 subtype and that several HPV DNA fragments amplified by PCR from genital neoplasia represent worldwide distributed variants of HPV68 and HPV70.

IT 175702-95-1

RL: PRP (Properties)

(nucleotide sequence of genome; sequence of genital human papillomavirus (HPV) types HPV68 and HPV70 related to potentially oncogenic HPV39)

L2 ANSWER 8 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1995:893000 CAPLUS

DOCUMENT NUMBER: 123:308176

09/885799

TITLE: Primers for detection of human papilloma virus
by nucleic acid amplification
INVENTOR(S): Meijer, Christophorus Joannes Lambertus Maria;
Van Den Brule, Adrianus Johannes Christiaan;
Walboomers, Jan Marcus Maria; Snijders, Petrus
Josephus Ferdinandus
PATENT ASSIGNEE(S): Stichting Researchfonds Pathologie, Neth.
SOURCE: PCT Int. Appl., 61 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9522626	A1	19950824	WO 1995-NL66	19950220
W: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UG				
RW: KE, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG				
CA 2183758	AA	19950824	CA 1995-2183758	19950220
AU 9516722	A1	19950904	AU 1995-16722	19950220
AU 685233	B2	19980115		
EP 746627	A1	19961211	EP 1995-908386	19950220
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE				
JP 09509062	T2	19970916	JP 1995-521732	19950220
US 6352825	B1	20020305	US 1996-696932	19961125
PRIORITY APPLN. INFO.:				
			EP 1994-200432	A 19940221
			EP 1994-202739	A 19940923
			WO 1995-NL66	W 19950220
AB	Oligonucleotide primers suitable for detection of genital human papilloma viruses by nucleic acid amplification e.g. PCR, NASBA, or LCR, are described for use in diagnosis. The primers are suitable for detecting many types of HPV and may include some base variability away from their 3'-ends and some alterations in the 3'-sequence provided the primer forms a perfect match in the 3'-region. Hybridization probes specific for amplification products.			
IT	165309-07-9			
	RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)			
	(nucleotide sequence, hybridization probe for HPV-58; primers for detection of human papilloma virus by nucleic acid amplification)			
L2	ANSWER 9 OF 14 CAPLUS COPYRIGHT 2002 ACS			
ACCESSION NUMBER:	1995:803803 CAPLUS			
DOCUMENT NUMBER:	123:307451			
TITLE:	Generation of entire human papillomavirus genomes by long PCR: frequency of errors produced during amplification			
AUTHOR(S):	Stewart, Ann-Charlotte M.; Gravitt, Patti E.; Cheng, Suzanne; Wheeler, Cosette M.			
CORPORATE SOURCE:	Dep. Cell Biol., Univ. New Mexico Cancer Res., Albuquerque, NM, 87131-5226, USA			

Searcher : Shears 308-4994

09/885799

SOURCE: Genome Res. (1995), 5(1), 79-88

CODEN: GEREFS

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Recently, several improvements of traditional PCR techniques have facilitated the amplification of significantly longer DNA target sequences. An improved method is reported for amplification of entire human papillomavirus (HPV) genomes. The rTth DNA polymerase, XL (Perkin-Elmer, Foster City, CA), and the accompanying XL PCR buffer system, were used to successfully amplify 8-kb genomes from .apprx.10 copies of input ref. strain HPV16 DNA. This long PCR (LPCR) method was subsequently used to amplify the entire HPV16 genome from clin. specimens. The fidelity with which the rTth DNA polymerase XL amplifies target sequences under the chosen amplification conditions was estd. by partial sequencing of cloned LPCR products generated from cloned ref. strain HPV16 genomes. A region spanning the HPV16 E6, E7, and part of the E1 open reading frames (ORFs) was sequenced in 29 clones. A total of 33 nucleotide substitutions were obsd. in the 23.5 kb sequenced. This corresponds to an error frequency of .apprx.1 error per 700 bases. Finally, LPCR methods were used to amplify entire, novel HPV genomes from clin. specimens. LPCR primer pairs were designed for amplification of 7 potentially novel HPV types. Amplicons of .apprx.8 kb were generated from 5 of the 7 HPV types targeted. One of the LPCR-generated novel genomes, CP141, was subsequently cloned and a partial sequence was detd.

IT 160074-98-6P, GenBank U12476

RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)

(nucleotide sequence; generation of entire human papillomavirus genomes by long PCR and amplification of novel HPV genomes)

L2 ANSWER 10 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1995:547519 CAPLUS

DOCUMENT NUMBER: 123:75766

TITLE: Group-specific differentiation between high- and low-risk human papillomavirus genotypes by general primer-mediated PCR and two cocktails of oligonucleotide probes

AUTHOR(S): Jacobs, M. V.; De Roda Husman, A. M.; van den Brule, A. J. C.; Snijders, P. J. F.; Meijer, C. J. L. M.; Walboomers, J. M. M.

CORPORATE SOURCE: Department Pathology, Free University Hospital, Amsterdam, 1081 HV, New Caledonia

SOURCE: J. Clin. Microbiol. (1995), 33(4), 901-5

CODEN: JCMIDW; ISSN: 0095-1137

DOCUMENT TYPE: Journal

LANGUAGE: English

AB In recent years, general primer-mediated PCR assays have been developed to detect a broad spectrum of human papillomavirus (HPV) genotypes. In this study, a procedure enabling a simple group-specific differentiation of high-risk (HPV-16, -18, -31, -33, -35, -39, -45, -51, -52, -54, -56, and -58) and low-risk (HPV-6, -11, -34, -40, -42, -43, and 44) HPVs following an HPV general primer-mediated (GP5+/GP6+) PCR is presented. By computer-assisted sequence anal., oligonucleotides (30-mers) specific for 19 different HPV genotypes were selected from the internal part of the 150-bp GP5+/GP6+-amplified region. These oligo probes were tested for

09/885799

specificity in a Southern blot anal. of PCR products derived from the same panel of HPV types. No cross-hybridizations were found. The sensitivities of the oligo probes varied from the femtogram level for the well-amplified HPV types like HPV-16 and -18 to the picogram level for the less-well-amplified HPV types like HPV-39 and -51. These sensitivities were reached when the oligo probes were applied both individually and in a cocktail. On the basis of these results, two cocktail oligo probes that enabled a specific and sensitive differentiation between low- and high-risk HPV types were composed.

IT 165309-07-9

RL: ARG (Analytical reagent use); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses) (human papillomavirus 58-specific probe; group-specific differentiation between high- and low-risk human papillomavirus genotypes by general primer-mediated PCR and two cocktails of oligonucleotide probes)

L2 ANSWER 11 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1995:225424 CAPLUS

DOCUMENT NUMBER: 122:283458

TITLE: Identification of genomic sequences of three novel human papillomavirus sequences in cervical smears of Amazonian Indians

AUTHOR(S): Ong, Chi-Keong; Bernard, Hans-Ulrich; Villa, Luisa L.

CORPORATE SOURCE: Institute of Molecular and Cell Biology, National University of Singapore, Singapore

SOURCE: J. Infect. Dis. (1994), 170(5), 1086-8

CODEN: JIDIAQ; ISSN: 0022-1899

DOCUMENT TYPE: Journal

LANGUAGE: English

AB During examn. of cervical smears from 180 American Indian patients from the Amazon, 3 of 26 samples that were pos. for human papillomavirus (HPV) infection by the polymerase chain reaction (PCR) did not hybridize with DNA probes to 19 different genital HPV types. The nucleotide sequences of these clones, LVX82, LVX100, and LVX160, confirmed their origin from HPV genomes. They were distinct from all presently known genital HPV types because each sequence differed by >10% compared with their closest relatives, HPV-39, HPV-61, and HPV-62. Thus, these three clones were derived from viruses that would qualify as new HPV types should their complete genomes be isolated. It was hoped that these clones would represent HPV types endemic to America's aboriginal populations; however, corresponding studies detected related HPV genomes in non-Indian populations of Brazil, North America, and Southeast Asia.

IT 160075-08-1, GenBank U12486

RL: PRP (Properties)

(nucleotide sequence of; genomic sequences of three novel human papillomavirus sequences in cervical smears of Amazonian Indians)

L2 ANSWER 12 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1995:212124 CAPLUS

DOCUMENT NUMBER: 122:232212

TITLE: Identification of five novel human papillomavirus sequences in the New Mexico triethnic population

AUTHOR(S): Peyton, Cheri L.; Wheeler, Cosette M.

Searcher : Shears 308-4994

09/885799

CORPORATE SOURCE: Cancer Research and Treatment Center, University
New Mexico, Albuquerque, NM, USA
SOURCE: J. Infect. Dis. (1994), 170(5), 1089-92
CODEN: JIDIAQ; ISSN: 0022-1899
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Cervical specimens obtained from previous clin. studies at the University of New Mexico were screened for ref. strain human papillomaviruses (HPVs) via HPV L1 consensus polymerase chain reaction (PCR) and type-specific oligonucleotide probe hybridization. PCR products that bound to an HPV generic probe but did not hybridize with HPV type-specific oligonucleotide probes were subjected to restriction endonuclease digestion analyses. Resulting restriction fragment length polymorphisms served as the basis for initial grouping of unknown HPVs and identification of potential ref. strain HPV variants. MY09/MY11 PCR products contg. putative novel HPV L1 fragments were cloned and sequenced. Five novel HPV sequences were identified in this study.

IT 160074-98-6, GenBank U12476

RL: PRP (Properties)
(nucleotide sequence; identification of five novel human papillomavirus sequences in New Mexico triethnic population)

L2 ANSWER 13 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1995:143122 CAPLUS

DOCUMENT NUMBER: 122:124831

TITLE: Analysis of short novel human papillomavirus sequences

AUTHOR(S): Tachezy, Ruth; Van Ranst, Marc A.; Cruz, Yvette; Burk, Robert D.

CORPORATE SOURCE: Dep. Pediatrics, Microbiol. Immunol., Albert Einstein Coll. of Med., Bronx, NY, 10461, USA

SOURCE: Biochem. Biophys. Res. Commun. (1994), 204(2), 820-7

CODEN: BBRCA9; ISSN: 0006-291X

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Consensus primer mediated PCR protocols have the potential to amplify previously uncharacterized human papillomavirus (HPV) genotypes. In a study on 500 cervicovaginal samples, we amplified four sequences (L1A# 1 to L1AE 4) that failed to hybridize to any of the available HPV type-specific oligonucleotide probes. Nucleotide sequencing revealed that the sequences were derived from the L1 region of hitherto unsequenced genotypes. Comparison of phylogenetic trees based on the amplified L1 sequences with E6-derived phylogenetic trees resulted in the identification of L1AE 1 and L1AE 2 as putative novel HPV PCR genotypes. L1AE 1 was related HPV 39, whereas L1AE 2 was related to HPV 51. The L1AE 3 and L1AE 4 sequences occupied L1-phylogenetic branches equiv. to the positions of HPV 66 and HPV 61, resp., in an E6-phylogenetic tree.

IT 152370-56-4

RL: PRP (Properties)
(nucleotide sequence and phylogenetic anal. of)

L2 ANSWER 14 OF 14 CAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1992:606045 CAPLUS

DOCUMENT NUMBER: 117:206045

TITLE: Human papillomavirus type 58 DNA sequence

Searcher : Shears 308-4994

09/885799

AUTHOR(S): Kirii, Yasuyuki; Iwamoto, Seiichi; Matsukura, Toshihiko
CORPORATE SOURCE: Kanebo Inst. Cancer Res., Osaka, 534, Japan
SOURCE: Virology (1991), 185(1), 424-7
CODEN: VIRLAX; ISSN: 0042-6822
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The complete nucleotide sequence of human papillomavirus type 58 (HPV 58) DNA cloned from an invasive cervical carcinoma was detd. The HPV 58 genome consists of 7824 nucleotides, contg. 37.9% of GC residues, and has a similar genome organization of other HPVs. On the nucleotide sequence level, it conserves the signal sequences for regulation of gene expression as with other genital HPVs and exhibits an extensive homol. with HPV 33 (77%). Comparative anal. of amino acid sequences reveals that HPV 58 is closely related to HPVs 16, 31, and 33, and is more distantly related to HPVs 6, 11, 18, and 39. HPVs 58, 16, 31, and 33 can be regarded as a group in HPV.

IT 140812-52-8, Deoxyribonucleic acid (human papilloma virus 58 7.82-kilobase fragment) 144197-89-7, Deoxyribonucleic acid (human papilloma virus 58 gene L1)
RL: PRP (Properties); BIOL (Biological study)
(nucleotide sequence of)

E1 THROUGH E16 ASSIGNED

L3 FILE 'REGISTRY' ENTERED AT 10:29:05 ON 04 APR 2002
16 SEA FILE=REGISTRY ABB=ON PLU=ON (160074-98-6/BI OR 165309-07-9/BI OR 175702-95-1/BI OR 140812-52-8/BI OR 144197-89-7/BI OR 152370-56-4/BI OR 160075-08-1/BI OR 164961-97-1/BI OR 174054-08-1/BI OR 174054-09-2/BI OR 174054-10-5/BI OR 174054-11-6/BI OR 174054-12-7/BI OR 195231-89-1/BI OR 250163-20-3/BI OR 250163-21-4/BI)

L3 ANSWER 1 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 250163-21-4 REGISTRY
CN 13: PN: US5981173 SEQID: 11 unclaimed DNA (9CI) (CA INDEX NAME)
SQL 1059
MF Unspecified
CI MAN

REFERENCE 1: 131:347519

L3 ANSWER 2 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 250163-20-3 REGISTRY
CN 12: PN: US5981173 SEQID: 10 unclaimed DNA (9CI) (CA INDEX NAME)
SQL 1521
MF Unspecified
CI MAN

REFERENCE 1: 131:347519

L3 ANSWER 3 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 195231-89-1 REGISTRY
CN 10: PN: WO0168915 SEQID: 11 unclaimed DNA (9CI) (CA INDEX NAME)
SQL 30
MF Unspecified
CI MAN

09/885799

REFERENCE 1: 135:268104

L3 ANSWER 4 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 175702-95-1 REGISTRY
CN DNA (human papillomavirus 70 gene L1 plus gene E6 plus gene E7 plus
flanks) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (human papilloma virus strain HPV70)
OTHER NAMES:
CN DNA (human papillomavirus 70 protein L1 gene plus E6 protein gene
plus E7 protein gene plus flanks)
CN DNA (human papillomavirus strain HPV70).
SQL 3283
MF Unspecified
CI MAN

REFERENCE 1: 131:347519

REFERENCE 2: 124:280570

L3 ANSWER 5 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 174054-12-7 REGISTRY
CN DNA (human papillomavirus 58 strain IS404 protein L 1
fragment-specifying) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (human papilloma virus 58 strain IS404 protein
L 1 fragment-specifying)
OTHER NAMES:
CN GenBank U45928
SQL 449
MF Unspecified
CI MAN

REFERENCE 1: 124:280734

L3 ANSWER 6 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 174054-11-6 REGISTRY
CN DNA (human papillomavirus 58 strain IS131 protein L 1
fragment-specifying) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (human papilloma virus 58 strain IS131 protein
L 1 fragment-specifying)
OTHER NAMES:
CN GenBank U45927
SQL 449
MF Unspecified
CI MAN

REFERENCE 1: 124:280734

L3 ANSWER 7 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 174054-10-5 REGISTRY
CN DNA (human papillomavirus 58 strain IS1021 protein L 1
fragment-specifying) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (human papilloma virus 58 strain IS1021
protein L 1 fragment-specifying)

09/885799

OTHER NAMES:

CN GenBank U45926
SQL 449
MF Unspecified
CI MAN

REFERENCE 1: 124:280734

L3 ANSWER 8 OF 16 REGISTRY COPYRIGHT 2002 ACS

RN 174054-09-2 REGISTRY

CN DNA (human papillomavirus 58 strain IS573 protein L 1
fragment-specifying) (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Deoxyribonucleic acid (human papilloma virus 58 strain IS573 protein
L 1 fragment-specifying)

OTHER NAMES:

CN GenBank U45925
SQL 449
MF Unspecified
CI MAN

REFERENCE 1: 124:280734

L3 ANSWER 9 OF 16 REGISTRY COPYRIGHT 2002 ACS

RN 174054-08-1 REGISTRY

CN DNA (human papillomavirus 58 strain IS068 protein L 1
fragment-specifying) (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Deoxyribonucleic acid (human papilloma virus 58 strain IS068 protein
L 1 fragment-specifying)

OTHER NAMES:

CN GenBank U45924
SQL 449
MF Unspecified
CI MAN

REFERENCE 1: 124:280734

L3 ANSWER 10 OF 16 REGISTRY COPYRIGHT 2002 ACS

RN 165309-07-9 REGISTRY

CN DNA, d(A-T-T-A-T-G-C-A-C-T-G-A-A-G-T-A-A-C-T-A-A-G-G-A-A-G-G-T-A-C)
(9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Deoxyribonucleic acid, d(A-T-T-A-T-G-C-A-C-T-G-A-A-G-T-A-A-C-T-A-A-G-
G-A-A-G-G-T-A-C)

SQL 30
MF Unspecified
CI MAN

REFERENCE 1: 134:217713

REFERENCE 2: 123:308176

REFERENCE 3: 123:75766

L3 ANSWER 11 OF 16 REGISTRY COPYRIGHT 2002 ACS

RN 164961-97-1 REGISTRY

CN DNA (human papillomavirus 70 strain KK57) (9CI) (CA INDEX NAME)

Searcher : Shears 308-4994

09/885799

OTHER CA INDEX NAMES:

CN Deoxyribonucleic acid (human papilloma virus 70 strain KK57)

OTHER NAMES:

CN GenBank U21941

SQL 7904

MF Unspecified

CI MAN

REFERENCE 1: 124:309077

L3 ANSWER 12 OF 16 REGISTRY COPYRIGHT 2002 ACS

RN 160075-08-1 REGISTRY

CN DNA (human papillomavirus strain LVX160 protein L 1 gene fragment)
(9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Deoxyribonucleic acid (human papilloma virus strain LVX160 protein L
1 gene fragment)

OTHER NAMES:

CN GenBank U12486

SQL 455

MF Unspecified

CI MAN

REFERENCE 1: 122:283458

L3 ANSWER 13 OF 16 REGISTRY COPYRIGHT 2002 ACS

RN 160074-98-6 REGISTRY

CN DNA (human papillomavirus strain CP141 protein L 1
fragment-specifying) (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Deoxyribonucleic acid (human papilloma virus strain CP141 protein L
1 fragment-specifying)

OTHER NAMES:

CN DNA (human papillomavirus strain CP141 protein L1
fragment-specifying)

CN GenBank U12476

SQL 455

MF Unspecified

CI MAN

REFERENCE 1: 124:252008

REFERENCE 2: 123:307451

REFERENCE 3: 122:232212

L3 ANSWER 14 OF 16 REGISTRY COPYRIGHT 2002 ACS

RN 152370-56-4 REGISTRY

CN DNA (human papillomavirus L1AE 4 type AE1 gene L1) (9CI) (CA INDEX
NAME)

OTHER CA INDEX NAMES:

CN Deoxyribonucleic acid (human papillomavirus L1AE 4 type AE1 gene L1)

OTHER NAMES:

CN GenBank U01535

SQL 415

MF Unspecified

CI MAN

09/885799

REFERENCE 1: 122:124831

L3 ANSWER 15 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 144197-89-7 REGISTRY
CN DNA (human papillomavirus 58 gene L1) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (human papilloma virus 58 gene L1)
SQL 1575
MF Unspecified
CI MAN

REFERENCE 1: 117:206045

ANSWER 16 OF 16 REGISTRY COPYRIGHT 2002 ACS
RN 140812-52-8 REGISTRY
CN DNA (human papillomavirus 58 7.82-kilobase fragment) (9CI) (CA
INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (human papilloma virus 58 7.82-kilobase
fragment)
SQL 7824
MF Unspecified
CI MAN

REFERENCE 1: 117:206045

FILE 'HOME' ENTERED AT 10:29:42 ON 04 APR 2002